

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2003, 10:00:02 ; Search time 3180 Seconds
(without alignments)
6998.376 Million cell updates/sec

Title: US-09-938-114-1
Perfect score: 544
Sequence: 1 GAAGGGCATGCTACAGGT.....ATTCTGCTAGCTGAAAAA 544

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_mam.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

ALIGNMENTS

RESULT 1	AR259035	544 bp	DNA	linear	PAT 20-DEC-2002
LOCUS	AR259035	Sequence 1 from patent US 6489451.			
DEFINITION	AR259035				
ACCESSION	AR259035				
VERSION	AR259035.1	GI:27309520			
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 544)				
AUTHORS	Li,B.X. and Cheng,X.				
TITLE	Antithrombin enzyme from the snake venom of agkistrodon acutus				
JOURNAL	Patent: US 6489451-A 1 03-DEC-2002;				
FEATURES	Location/Qualifiers				

Pred. No. is the number of results predicted by chance to have a

source	1. 544		
BASE COUNT	158 a	122 c	143 g
ORIGIN	/organism="unknown"		
Query Match	100.0%;	Score 544;	DB 6; Length 544;
Best Local Similarity	100.0%;	Pred. No. 1.2e-151;	
Matches 544;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	GAAGGGCATTGCTACAAAGTCTTCAAAACATCTAAGACCTTGGACAGATGACAGAGCTTC	60
Db	1	GAAGGGCATTGCTACAAAGTCTTCAAAACATCTAAGACCTTGGACAGATGACAGAGCTTC	60
QY	61	TGCACGAAGCAGGTGAACGGGGGCGCATCTGCTCTATCGAAAGCTCCGGAGAGCAGAC	120
Db	61	TGCACGAAGCAGGTGAACGGGGGCGCATCTGCTCTATCGAAAGCTCCGGAGAGCAGAC	120
QY	121	TTTGTGGCCAGTTGATGCTCAGAGATTAAGTTCAGCCAAATCCATGCTCGATCGGA	180
Db	121	TTTGTGGCCAGTTGATGCTCAGAGATTAAGTTCAGCCAAATCCATGCTCGATCGGA	180
QY	181	CTGAGGGCTCAAAACAAAGAAAGCAATGCAGATAGAGTGGAGCGATGCTCCAGCATC	240
Db	181	CTGAGGGCTCAAAACAAAGAAAGCAATGCAGATAGAGTGGAGCGATGCTCCAGCATC	240
QY	241	AGTTATGAGAAATTGGAATGAAGAAGAAATCCAAAAGTGTCTTGGGTGCACATAGAGACA	300
Db	241	AGTTATGAGAAATTGGAATGAAGAAGAAATCCAAAAGTGTCTTGGGTGCACATAGAGACA	300
QY	301	GGGTTTCATAAGTGGGAGAAATTTTACTGTGAACAAAGATCTTTTGTCTGCGAGGCA	360
Db	301	GGGTTTCATAAGTGGGAGAAATTTTACTGTGAACAAAGATCTTTTGTCTGCGAGGCA	360
QY	361	TAGTCTGAAGATCCAGCTGATTAAGTCTGGAAGCAAGAGAGCCGCCACCCATCCC	420
Db	361	TAGTCTGAAGATCCAGCTGATTAAGTCTGGAAGCAAGAGAGCCGCCACCCATCCC	420
QY	421	CCAAACCTTCCTAGCCACAAATCTTGCTATGACCCCTTTGCTCAAGGATGCTCTCTGTA	480
Db	421	CCAAACCTTCCTAGCCACAAATCTTGCTATGACCCCTTTGCTCAAGGATGCTCTCTGTA	480
QY	481	GCTGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGTA	540
Db	481	GCTGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGTA	540
QY	541	AAAA 544	
Db	541	AAAA 544	
RESULT 2			
AF176420			
LOCUS			
DEFINITION	Deinagkistrodon acutus clone 2100490 agkiscutacin A chain mRNA, complete cds.		
ACCESSION	AF176420 AY091757		
VERSION	AF176420.2 GI:20273043		
KEYWORDS	Deinagkistrodon acutus		
SOURCE	Deinagkistrodon acutus		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea; Viperidae; Crotalinae; Deinagkistrodon.		
REFERENCE	1 (bases 1 to 733)		
AUTHORS	Yu, H.-X., Xiang, K.-J., and Liu, J.		
TITLE	cDNA sequencing and analysis of eleven C-type lectin-like protein subunits from Agkistrodon acutus		
JOURNAL	Sheng wu hua hseuh yu sheng wu li hseuh pao (2002) In press		
REFERENCE	2 (bases 1 to 733)		
AUTHORS	Cheng, X., Qian, Y., Liu, Q., Li, B. X. Y., Ding, J., Xu, Z., Huang, W. and Liu, J.		
TITLE	Agkiscutacin, a new fibrinolytic & anti-platelet protein from Agkistrodon acutus venom		

JOURNAL	Unpublished			
REFERENCE	3 (bases 1 to 733)			
AUTHORS	Yu, H.-X., Xiang, K.-J., Wang, Y. and Liu, J.			
TITLE	A chain of agkiscutacin from Deinagkistrodon acutus			
JOURNAL	Unpublished			
REFERENCE	4 (bases 1 to 733)			
AUTHORS	Cheng, X., Liu, J., Li, B. X. Y. and Qian, Y.			
TITLE	Direct Submission			
JOURNAL	Submitted (08-AUG-1999) Dept. of Biochemistry and Molecular Biology, University of Science and Technology of China, School of Life-Science, Huangshan Road, Hefei, Anhui 230027, China			
REFERENCE	5 (bases 1 to 733)			
AUTHORS	Yu, H.-X., Xiang, K.-J., Wang, Y. and Liu, J.			
TITLE	Direct Submission			
JOURNAL	Submitted (23-APR-2002) Dept. of Biochemistry and Molecular Biology, University of Science and Technology of China, School of Life-Science, Huangshan Road, Hefei, Anhui 230027, China			
REMARK	Sequence update by submitter			
COMMENT	On Apr 23, 2002 this sequence version replaced gi:6715112.			
FEATURES	Location/Qualifiers			
Source	1. 733			
	/organism="Deinagkistrodon acutus"			
	/mol_type="mRNA"			
	/db_xref="taxon:36307"			
	/clone="2100490"			
	/tissue_type="venom"			
	89. 547			
CDS	/note="consists of two heterologous subunits (A and B chains) linked by an inter-subunit disulfide bond; c-type lectin family member"			
	/codon_start=1			
	/product="agkiscutacin A chain"			
	/protein_id="AAF26286.2"			
	/db_xref="GI:20273044"			
BASE COUNT	191 a 173 c 204 g 165 t			
ORIGIN				
Query Match	93.2%;	Score 507.2;	DB 5;	Length 733;
Best Local Similarity	98.0%;	Pred. No. 1.3e-140;		
Matches 535;	Conservative 0;	Mismatches 8;	Indels 3;	Gaps 2;
QY	1	GAAGGGCATTGCTACAAAGTCTTCAAAACATCTAAGACCTTGGACAGATGACAGAGCTTC	60	
Db	185	GAAGGGCATTGCTACAAAGTCTTCAAAACATCTAAGACCTTGGACAGATGACAGAGCTTC	244	
QY	61	TGCACGAAGCAGGTGAACGGGGGCGCATCTGCTCTATCGAAAGCTCCGGAGAGCAGAC	120	
Db	245	TGCACGAAGCAGGTGAACGGGGGCGCATCTGCTCTATCGAAAGCTCCGGAGAGCAGAC	304	
QY	121	TTTGTGGCCAGTTGATGCTCAGAAGATAAGTCAGCCAAATCCATCTCTGGATCGGA	180	
Db	305	TTTGTGGCCAGTTGATGCTCAGAAGATAAGTCAGCCAAATCCATCTCTGGATCGGA	364	
QY	181	CTGAGGGCTCAAAACAAAGAAAGCAATGCAGATAGAGTGGAGCGATGGCTCCAGCATC	240	
Db	365	CTGAGGGCTCAAAACAAAGAAAGCAATGCAGATAGAGTGGAGCGATGGCTCCAGCATC	424	
QY	241	AGTTATGAGAAATTGGAATGAAGAAGAAATCCAAAAGTGTCTTGGGTGCACATAGAGACA	300	
Db	425	AGTTATGAGAAATTGGAATGAAGAAGAAATCCAAAAGTGTCTTGGGTGCACATAGAGACA	484	
QY	301	GGGTTTCATAAGTGGGAGAAATTTTACTGTGAACAAAGATCCCTTTTGTCTGCGAGGCA	360	
Db	485	GGGTTTCATAAGTGGGAGAAATTTTACTGTGAACAAAGATCCCTTTTGTCTGCGAGGCA	544	
QY	361	TAGTCTGAAGATCCAGCTGATTAAGTCTGGAAGCAAGAGAGCCGCCACCCATCCC	420	
Db	545	TAGTCTGAAGATCCAGCTGATTAAGTCTGGAAGCAAGAGAGCCGCCACCCATCCC	604	
QY	421	CCAACCCCTCCTAGCCACAAATCTGCTATGACCCCTTTGCTCAAGGATGCTCTCTGTA	480	

206 TGCACGAAGCGGTGACGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAAGCAGAC

Matches 384; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 GAAGGGCATTCCTACAGGCTCTCAACAATCTAAGACCTGGACAGATGACAGAGCTTC 60
 |||||
 Db 97 GAAGGGCATTCCTACAGGCTCTCAACAATCTAAGACCTGGACAGATGACAGAGCTTC 156
 |||||

QY 61 TGCACGAGCAGGTGAACGGGGGCGCATCTGTCTCTATCGAAAGCTCCGGAGAGCAGAC 120
 |||||
 Db 157 TGCACGAGCAGGTGAACGGGGGCGCATCTGTCTCTATCGAAAGCTCCGGAGAGCAGAC 216
 |||||

QY 121 TTGTGTGGCCAGTGTGATCTCAGAGATAAAGTCAAGCAAAATCCATGTCTGGATCGGA 180
 |||||
 Db 217 TTGTGTGGCCAGTGTGATCTCAGAGATAAAGTCAAGCAAAATCCATGTCTGGATCGGA 276
 |||||

QY 181 CTGAGGGCTCAAAACAAAGAAAGCAATGAGATAGAGTGGACCGATGCTCCAGCATC 240
 |||||
 Db 277 CTGAGGGCTCAAAACAAAGAAAGCAATGAGATAGAGTGGACCGATGCTCCAGCATC 336
 |||||

QY 241 AGTTATGAGAAATGGATTGAAGAAAGTCAAAAGTGTCTTGGGGTGACATAGAGACA 300
 |||||
 Db 337 AGTTATGAGAAATGGATTGAAGAAAGTCAAAAGTGTCTTGGGGTGACATAGAGACA 396
 |||||

QY 301 GGGTTTCATAAGTGGGAGAAATTTTACTGTGAACAACAGATCCTTTGTCTCGAGGCA 360
 |||||
 Db 397 GGGTTTCATAAGTGGGAGAAATTTTACTGTGAACAACAGATCCTTTGTCTCGAGGCA 456
 |||||

QY 361 TAGTCTGAAGATCCAGCTGATGAAGTCTGGAGAA 395
 |||||
 Db 457 TAGTCTGAAGATCCAGCTGATGAAGTCTGGAGAA 491
 |||||

RESULT 5
 LOCUS
 DEFINITION
 Trimeresurus flavoviridis mRNA for factor IX/factor X binding
 protein A chain, complete cds.

ACCESSION
 D83331.1

VERSION
 D83331.1 GI:1402639

KEYWORDS
 factor IX/factor X binding protein A chain.

SOURCE
 Trimeresurus flavoviridis
 Trimeresurus flavoviridis
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
 Viperidae; Crotalinae; Trimeresurus.
 1 (bases 1 to 678)
 Matsuzaki, R., Yoshiara, E., Yamada, M., Shima, K., Atoda, H. and
 Morita, T.
 cDNA cloning of IX-BP, a heterogeneous two-chain anticoagulant
 protein from snake venom
 Biochem. Biophys. Res. Commun. 220 (2), 382-387 (1996)

REFERENCE
 8645314
 Morita, T.
 Direct Submission
 Submitted (02-FEB-1996) Takashi Morita, Meiji College of Pharmacy,
 Biochemistry; 1-22-1 Yatocho, Tanashi, Tokyo 188, Japan
 (Tel:0424-21-0101(ex.429), Fax:0424-21-1489)
 Location/Qualifiers
 1..678
 /organism="Trimeresurus flavoviridis"
 /mol_type="mRNA"
 /db_xref="taxon:88087"
 1..49
 50..508
 /product="factor IX/factor X binding protein A chain"
 /protein_id="BA011887.1"
 /db_xref="GI:1402640"
 /translation="MGFIFMFGFLIVVAASLRGTGADCLSGWSSVGHGCHYKAFKVK
 TWEDAEVCTEQAKGLVLSIESGADFVAQLVTQNMKELDFYIWIQLVQKVKQC
 NSEWSDSSVSVENWIEAESKTCIGLEKETDFRKNYNYICGQONPFVCEA"
 509..678

3' UTR

polyA_signal 659..664
 BASE COUNT 174 a 161 c 191 g 152 t
 ORIGIN

Query Match 69.2%; Score 376.2; DB 5; Length 678;
 Best Local Similarity 83.7%; Pred. No. 2.3e-101;
 Matches 453; Conservative 0; Mismatches 78; Indels 10; Gaps 2;

QY 1 GAAGGGCATTCCTACAGGCTCTCAACAATCTAAGACCTGGACAGATGACAGAGCTTC 60
 |||||
 Db 146 GAAGGGCATTCCTACAGGCTCTCAACAATCTAAGACCTGGAGAGATGACAGAGCTTC 205
 |||||

QY 61 TGCACGAGCAGGTGAACGGGGGCGCATCTGTCTCTATCGAAAGCTCCGGAGAGCAGAC 120
 |||||
 Db 206 TGCACGAGCAGGTGAACGGGGGCGCATCTGTCTCTATCGAAAGCTCCGGAGAGCAGAC 265
 |||||

QY 121 TTGTGTGGCCAGTGTGATCTCAGAGATAAAGTCAAGCAAAATCCATGTCTGGATCGGA 180
 |||||
 Db 266 TTGTGTGGCCAGTGTGATCTCAGAGATAAAGTCAAGCAAAATCCATGTCTGGATCGGA 325
 |||||

QY 181 CTGAGGGCTCAAAACAAAGAAAGCAATGAGATAGAGTGGACCGATGCTCCAGCATC 240
 |||||
 Db 326 CTGAGGGCTCAAAACAAAGAAAGCAATGAGATAGAGTGGACCGATGCTCCAGCATC 385
 |||||

QY 241 AGTTATGAGAAATGGATTGAAGAAAGTCAAAAGTGTCTTGGGGTGACATAGAGACA 300
 |||||
 Db 386 AGTTATGAGAAATGGATTGAAGAAAGTCAAAAGTGTCTTGGGGTGACATAGAGACA 445
 |||||

QY 301 GGGTTTCATAAGTGGGAGAAATTTTACTGTGAACAACAGATCCTTTGTCTCGAGGCA 360
 |||||
 Db 446 GATTTTCGTAAGTGGGTCAATATTTACTGTGGACAAACAAATCTCTGCTGCGAGGCA 505
 |||||

QY 361 TAGTCTGAAGATCCAGCTGATGAAGTCTGGAGAAAGCAAGAGAGAGAGAGAGAGAGAG 420
 |||||
 Db 506 TAGTCTGAAGATCCAGCTGATGAAGTCTGGAGAAAGCAAGAGAGAGAGAGAGAGAG 557
 |||||

QY 421 CCAACCCCTGCTGAGCCCAATCTCTGCTATGCACCTTTGCTCAACGGATGCTCTCTGTA 480
 |||||
 Db 558 CCAACCCCTGCTGAGCCCAATCTCTGCTATGCACCTTTGCTCAACGGATGCTCTCTGTA 617
 |||||

QY 481 GCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 538
 |||||
 Db 618 GCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 677
 |||||

QY 539 G 539
 Db 678 G 678

RESULT 6
 LOCUS
 DEFINITION
 Trimeresurus flavoviridis mRNA for factor XI/factor X binding
 protein A chain, complete cds.

ACCESSION
 AB046491

VERSION
 AB046491.1

KEYWORDS
 GI:12583676

SOURCE
 Trimeresurus flavoviridis
 Trimeresurus flavoviridis
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
 Viperidae; Crotalinae; Trimeresurus.
 1 (sites)
 Tani, A., Ogawa, T., Nose, T., Nikandorov, N.N., Deshimaru, M.,
 Chijiwa, T., Chang, C., Fukumaki, Y. and Ohno, M.
 Characterization and molecular evolution of an anticoagulant
 protein from Agkistrodon actus venom
 Unpublished
 2 (bases 1 to 683)
 Ogawa, T. and Ohno, M.
 Direct Submission
 Submitted (24-JUL-2000) Tomohisa Ogawa, Tohoku University, Dept. of
 Agricultural Science; 1-1 Teutsuimidori Amamiyamachi, Aoba-ku,

Sendai, Miyagi 981-8555, Japan (E-mail:ogawa@biochem.tohoku.ac.jp,
Tel:81-22-717-8808, Fax:81-22-717-8807)

```

FEATURES
  Source
    1..683
      /organism="Trimeresurus flavoviridis"
      /mol_type="mRNA"
      /db_xref="taxon:88087"
      /tissue_type="venome gland"
      /country="Japan:Kagoshima, Tokunoshima island"
  gene
    1..683
      /gene="XI/Xbp-a"
  CDS
    77..517
      /gene="XI/Xbp-a"
      /codon_start=1
      /product="factor XI/factor X binding protein A chain"
      /protein_id="BAB21452.1"
      /db_xref="GI:12583677"
      /translation="MSFGLLVVAASLRGTGADCLSGSSSYGHCHCYKAFKVKYTWEDAE
RVCTEQAGHLVSTESGEADFAQLVITQNMRLDFYIWLIGRVQGVKQCNSEWSD
GSSVSYENWIEAESKICLGEKETDFRKNVNIYCGQNPFFCEA"
  BASE COUNT
    172 a 162 c 194 g 155 t
  ORIGIN
    65..241; Score 354.8; DB 5; Length 683;
    Best Local Similarity 82.7%; Pred. No. 6e-95;
    Matches 445; Conservative 0; Mismatches 82; Indels 11; Gaps 3;
  QY 1 GAAGGGCATTGCTACAAAGTCTTCAAAACATCTAAGACCTGCACAGATGCAGAGAGCTTC 60
  Db 155 GAGGGCATTGCTACAAAGCCTTCGAAATATACAGACCTGGAGATGCAGAGAGCTTC 214
  QY 61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAGCAGAC 120
  Db 215 TGCACGGAGCAGCGGAGGGCGGCATCTGGTCTCTATCGAAAGCTCCGGAGAGCAGAC 274
  QY 121 TTTGTGGCCAGTTGATTCCTCAGAGATAAGTACGCCAAATCCATCTCTGGATCGGA 180
  Db 275 TTTGTGGCCAGCTGGTCTACTCAGAACATGAAGAGACTCGATTTCTATATCTGGATCGGG 334
  QY 181 CTGAGGGCTCAAAACAAAGAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
  Db 335 CTGAGGGTTCAAGGAAGTCAAGCAATGCAGTGGAGTGGAGCGATGGCTCCAGCGTC 394
  QY 241 AGTTATGAGAATTGGAATGAAGAAGATCCAAAAGCTGTCTTTGGGTGCACATAGAGACA 300
  Db 395 AGTTATGAGAATTGGAATGAAGAAGATCCAAAAGCTGTCTTTGGGTGCACATAGAGACA 454
  QY 301 GGGTTTCATAGTGGAGAAATTTTACTGTGACACAAAGATCCCTTTTCTGCGAGGCA 360
  Db 455 GATTTTCGTAAGTGGGTCAATATTTACTGTGGAACAAAGATCCCTTTTCTGCGAGGCA 514
  QY 361 TAGTCTGAAGATCCAGTCAATGAAGTCTGGAGAGCAAGAGAGAGAGAGAGAGAGAGAG 420
  Db 515 TAGTCTGAAGATCCAGTCTGAGTCTGAGTCTGGAGAGCAAGAGAGAGAGAGAGAGAGAG 566
  QY 421 CCACCCCTGCTAGCCACAAATCTCTGCTATGCAACCTTTGCTCAACGAGATGCTCTCTGA 480
  Db 567 CAACACCCACCTGCGCCCAATCTGCTCTGCAACC--TTGCTCAACGAGATGCTCTCTGA 625
  QY 481 GCTGGATCTGGTGTCTGCTGCTGATGGCCGGAG--TCATATAATCTGCTAGC 536
  Db 626 GCTGGATCTGGTGTCTGCTGCTGATGGCCGGAGAGTCCAATAATCTGCTAGC 683

```

RESULT 7
AF540646
LOCUS
DEFINITION
Deinagkistrodon acutus agglucetin-alpha 2 subunit precursor, mRNA,
complete cds.
ACCESSION
AF540646
VERSION
AF540646.1 GI:23321260
KEYWORDS
Deinagkistrodon acutus

```

ORGANISM
  Deinagkistrodon acutus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;
  Viperidae; Crotalinae; Deinagkistrodon.
  1 (bases 1 to 661)
  Wang, W.-J. and Huang, T.-F.
  A novel tetrameric venom protein, agglucetin from Agkistrodon
  acutus, acts as a glycoprotein Ib agonist
  Thromb. Haemost. 86 (4), 1077-1086 (2001)
  21542539
  MEDLINE
  11686327
  REFERENCE
  2 (bases 1 to 661)
  Wang, W.-J., Ling, Q.-D. and Huang, T.-F.
  Molecular structure and functional characterization of agglucetin,
  a tetrameric glycoprotein Ib-binding protein, from Formosan pit
  viper
  Unpublished
  3 (bases 1 to 661)
  Wang, W.-J., Ling, Q.-D. and Huang, T.-F.
  Direct Submission
  Submitted (23-AUG-2002) Department of Nursing, Chang Gung Institute
  of Technology, 261, Wen-Hua 1 Rd., Kwei-Shan, Tao-Yuan, Taiwan
  333-03, R.O.C.
  FEATURES
    Location/Qualifiers
      1..661
        /organism="Deinagkistrodon acutus"
        /mol_type="mRNA"
        /db_xref="taxon:36307"
      13..489
        /note="platelet agglutination inducer; glycoprotein Ib
        agonist"
        /codon_start=1
        /product="agglucetin-alpha 2 subunit precursor"
        /protein_id="AAN23125.1"
        /db_xref="GI:23321261"
        /translation="MGRIFFVFLSLGTGADFNCPGWSAYDQYCYVKE
        KPNWDAERFCTEQADGHLVSIKSGRDFVAQLVSONIESVDEHVTGLRVONKEK
        QCSWSDSGSSVSLENLLELYMRKCALERETGFHKWNLGSCIQLNPFVCKPPQC"
      13..486
        sig_peptide
        mat_peptide
        /product="agglucetin-alpha 2 subunit"
  BASE COUNT
    167 a 159 c 183 g 152 t
  ORIGIN

```

```

  Query Match 61.7%; Score 335.8; DB 5; Length 661;
  Best Local Similarity 79.6%; Pred. No. 2.9e-89;
  Matches 437; Conservative 0; Mismatches 102; Indels 10; Gaps 3;
  QY 1 GAAGGCATTGCTACAAAGTCTTCAAAACATCTAAGACCTGGACAGATGCAGAGAGCTTC 60
  Db 115 GATCATGATGCTTACAGGTCTATCAAGAACCGAAACTGGATGATGCAGAGCGGTTC 174
  QY 61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAGAGAGAC 120
  Db 175 TGCACGGAGCAGCGGAGCGGTGGCCACCTGGTCTCTATCGAAAGCAAGGAGAGAGAGAC 234
  QY 121 TTTGTGGCCAGTTGATTCCTCAGAGATAAGTACGCCAAATCCATCTCTGGATCGGA 180
  Db 235 TTTGTGGCCAGCTGGTCTCTCAGAACATAGAGAGTGTTCAGAGACCATGTCTGGACCGGA 294
  QY 181 CTGAGGGCTCAAAACAAAGAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
  Db 295 CTGAGGGTTCAAAACAAAGAAAGCAATGCAGCAGATGGAGCGATGGCTCCAGCGTC 354
  QY 241 AGTTATGAGAATTGGAATGAAGAAGATCCAAAAGTGTCTTTGGGTGCACATAGAGACA 300
  Db 355 AGTTATGAGAATTGGAATGAAGAAGATCCAAAAGTGTCTTTGGGTGCACATAGAGACA 414
  QY 301 GGGTTTCATAGTGGAGAAATTTTACTGTGAACAAACAAGATCCCTTTTGTCTCGAG--- 357
  Db 415 GGGTTTCATAGTGGAGTCAATCTGGCTGTATACAACTGAATCCTTTTGTCTCGAGTTC 474
  QY 358 ---GCATAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAGCAAGGAGAGAGAGAGAGAG 414

```

```

Db      475  CCACCAAGTGTAAAGATCCAGCTGAGTGAAGTCTGAGAGCAAGAAAGCCCCACCA 534
QY      415  CATCCCCCAACCTGCTAGCACAATCTCTGCTATGCAACCTTTGCTCAACGATGCTC 474
Db      535  CA--GCCACCCCTGCTGCTCCGCAATCTCTGCTCTGACCCCTTTGCTCAATGATGCTC 592
QY      475  TCTGTAGCTGGATCTGCTGTGTTGCTGCTCCTGATGGGCGGAAG--TCAATAAATTTCTGCC 532
Db      593  TCTGTAGCTGGATCTGCTGTGTTGCTGCTCCTGATGGGCGGAAGTCAATAAATTTCTGCC 652
QY      533  TAGCCTGAA 541
Db      653  TAGCAAAA 661

RESULT 8
AF190827
LOCUS   Gloydius halys halysin A-chain precursor (HXNA) mRNA linear VRT 01-OCT-2000
DEFINITION
ACCESSION AF190827
VERSION AF190827.1 GI:10441755
KEYWORDS
SOURCE Gloydius halys (halys viper)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Lepidodonta; Squamata; Scleroglossa; Serpentes; Colubroidea;
Viperidae; Crotalinae; Gloydius.
REFERENCE 1 (bases 1 to 630)
AUTHORS Koo,B.H., Sohn,Y.D., Kim,D.S., Jang,Y.S. and Chung,K.H.
TITLE A novel coagulation factor Xa inhibitor from Korean snake
(Jagkistrodon halys) venom
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 630)
AUTHORS Koo,B.H., Sohn,Y.D., Kim,D.S., Jang,Y.S. and Chung,K.H.
TITLE Direct Submission
JOURNAL Submitted (29-SEP-1999) Cardiovascular Center, Yonsei University
College of Medicine, 134 Shinchon-dong, Seoul 120-752, Republic of
Korea

FEATURES
Source
1..630
/organism="Gloydius halys"
/mol_type="mRNA"
/db_xref="taxon:8714"
1..630
/gene="HXNA"
24..482
/gene="HXNA"
/function="coagulation factor Xa inhibitor"
/product="halyxin A-chain precursor"
/protein_id="AAG17178.1"
/db_xref="GI:10441756"
/translation="MGRIPLFSLGVLVFLSLGTGADCPGWSYEGHCYNIPHLFK
TWAEKRCQVKGHLVLESSEEDFVQLVSENKRYGIWIGLIRYGRKKQC
SSQWSDGSSVSYQNWIEASKTCLGQKTEFRKWNFYCGERNPFVCEA"
24..92
/gene="HXNA"
93..479
/gene="HXNA"
/product="halyxin A-chain"
173 a 137 c 174 g 146 t

sig_peptide
mat_peptide

BASE COUNT 173 a 137 c 174 g 146 t
ORIGIN

Query Match 57.6%; Score 313.4; DB 5; Length 630;
Best Local Similarity 84.2%; Pred. No. 1.5e-82;
Matches 353; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 1 GAAGGGCATTCCTCAAGTCTTCAACAATCTAAGACCTGACAGATGACAGAGCTTC 60
Db 120 GAAGGGCATTCCTCAACATCTTCATCTAATCAAGACCTGGCAGAGCAGAGGTTTC 179
QY 61 TGCAGAGCAGGTGAACGGGGGCGCATCTGCTCTCTATCGAAAGCTCGGAGAGCAGAC 120

```

```

Db      180  TGCAGAGCAGGTGAAGGGCGCGCATCTGCTCTATCGAAAGCTCCGAGAGCAGAC 239
QY      121  TTTGTGGCCAGTTGCTCAGAAAGATAAAGTCAAGCCAAATCCATGTCGATCGGA 180
Db      240  TTTGTGGCCAGTGTCTCTGAGACATGAAGAGATACGGGATCTATATCTCGATCGGA 299
QY      181  CTGAGGCTCAAAACAAAGAAAGCAATGAGTGCAGATAGAGTGGAGCGATGCTCCAGCATC 240
Db      300  CTGAGGCTTCGAGGCAAAAGCAATGAGTGCAGCTCCAGTGGAGCGATGCTCCAGGCTC 359
QY      241  AGTTATGAGATTCGATGAGAGAGATCCAAAAGTCTTGGGGTGCACATAGAGACA 300
Db      360  AGTTATGAGATTCGATGAGAGAGATCCAAAAGTCTTGGGGTGCACATAGAGACA 419
QY      301  GGGTTTCATAAGTGGAGAGATTTTACTGTGAACAACAAGATCTTTTGTCTCGAGGCA 360
Db      420  GAGTTTCGTAAGTGTATTAATTTACTGTGAGAGAGAAATCTTCTGCTCGAGGCA 479
QY      361  TAGTCTGAGATCCAGTGAATGAGTCTGGAGAAAGCAAGAGCCGCCCATCC 419
Db      480  TAGTCTGAGATCCAGTGTGAGTCTGGAGAAAGCAAGAGCCGCCCATCC 538

RESULT 9
AF540645
LOCUS   Deinagkistrodon acutus agglucetin-alpha 1 subunit precursor, mRNA,
DEFINITION complete cds.
ACCESSION AF540645
VERSION AF540645.1 GI:23321258
KEYWORDS
SOURCE Deinagkistrodon acutus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Lepidodonta; Squamata; Scleroglossa; Serpentes; Colubroidea;
Viperidae; Crotalinae; Deinagkistrodon.
REFERENCE 1 (bases 1 to 666)
AUTHORS Wang,W.-J. and Huang,T.F.
TITLE A novel tetrameric venom protein, agglucetin from Agkistrodon
acutus, acts as a glycoprotein Ib agonist
JOURNAL Thromb. Haemost. 86 (4), 1077-1086 (2001)
MEDLINE 21542539
PUBMED 11686327
REFERENCE 2 (bases 1 to 666)
AUTHORS Wang,W.-J., Ling,Q.-D. and Huang,T.-F.
TITLE Molecular structure and functional characterization of agglucetin,
a tetrameric glycoprotein Ib-binding protein, from Formosan pit
viper
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 666)
AUTHORS Wang,W.-J., Ling,Q.-D. and Huang,T.-F.
TITLE Direct Submission
JOURNAL Submitted (23-AUG-2002) Department of Nursing, Chang Gung Institute
of Technology, 261, Wen-Hua 1 Rd., Kwei-Shan, Tao-Yuan, Taiwan
333-03, R.O.C.

FEATURES
Location/Qualifiers
1..666
/organism="Deinagkistrodon acutus"
/mol_type="mRNA"
/db_xref="taxon:36307"
18..482
/feature="platelet agglutination inducer; glycoprotein Ib
agonist"
/codon_start=1
/product="agglucetin-alpha 1 subunit precursor"
/protein_id="AAN23124.1"
/db_xref="GI:23321258"
/translation="MGRIPLFSLGVLVFLSLGTGADVDCLPGWSYDQSCYRVFKL
LKWDDAEKCTERPKGHLVLESSEEDFVQLVSENKRYGIWIGLIRYGRKKQC
STWTDGSSVSYQNWIEASKTCLGQKTEFRKWNFYCGERNPFVCEA"
18..86
sig_peptide
87..479
mat_peptide

```

```
/product="agglucetin-alpha 1 subunit"
BASE COUNT      166 a   163 c   178 g   159 t
ORIGIN
Query Match      57.4%; Score 312; DB 5; Length 666;
Best Local Similarity 78.7%; Pred. No. 4e-82;
Matches 428; Conservative 0; Mismatches 100; Indels 16; Gaps 4;
QY 9 TTGCTACAAAGCTCTTCAAAACAACTTAAGACCTGGACAGATGCAGAGCTTCTGCACGAA 68
Db 128 TTGCTACCGGCTCTCAAACTCTCAAGCTCTGGATGATGCAGAGAGTTCTGCACGGA 187
QY 69 GCAGGTGAACCGGGGGCATCTGGTCTCTATCGAAAGTCCGGAGAGCAGACTTTTGTGGG 128
Db 188 GCGGCCGAAAGCGGGCATCTGGTCTCTATCGAAAGCGGGAGAGAGACTTTTGTGGC 247
QY 129 CCAGTTGATTGCTCAGAAGATAAGTCAAGCCAAATCATCTGATCGGACTGAGGSC 188
Db 248 CCAGCTGGTCTCTGAGAAACAAACAGACGACA-----ATGTCCTGGCTTGGACTGAAGAT 301
QY 189 TCAAACAAAGAAAGCAATGAGCATAGAGTGGAGCGATGGCTCCAGCATCAGTTATGA 248
Db 302 TCAAGCAAGAGCAGCAATGCAGACAGAGTGGACCGATGGCTCCAGGTCAGTTATGA 361
QY 249 GAATGGATTGAAGAGAAATCCAAAAGTGCTTGGGGTGCAATAGAGACAGGGTTTCA 308
Db 362 GAATTTTCTGAATACCAATCCAAAAGTGTTTGTGTGGAAGAAACACAGGGTTTGC 421
QY 309 TAAGTGGAGATTTTACTGTGAACACAGATCCTTTTGTCTGCA-----GGCATA 362
Db 422 CACATGGCTCAATCTTAATTTGGGAAGTGAATATGCTTTCGTCGAGATCCCGCCATA 481
QY 363 GTCTGAAGATCAGCTGATTTGAAGTCTGGAGAGCAAGGAAGCCCGCCACCCATCCCCC 422
Db 482 GTGTTAAGATCAGCTGAGTGAAGTCTGGAGAGCAAGGAAGCCCGCCACCCCA--GCC 539
QY 423 AACCTGCTAGCCACAATCTCTGCTATGCATGCAACCTTTTGTCTAAGGATGCTCTGTAGC 482
Db 540 ACCCCCTGCTGCGGCTATCTGCTCTGCAACCTTTTGTCTCAACGATGCTCTCTGTAGC 599
QY 483 TGGATCTGTTGTTGCTGCTCTGATGGCGCGGAG--TCAATAAATCTGCTAGCTGCTGA 540
Db 600 TGGATCTGTTTGTGCTCTGCTGATGGCGCGGAGTCCAAATAAATCTGCTAGCAAAA 659
QY 541 AAAA 544
Db 660 AAAA 663

RESULT 10
AX427204
LOCUS      AX427204
DEFINITION Sequence 4 from Patent WO0214364.
ACCESSION AX427204
VERSION    AX427204.1 GI:21530566
KEYWORDS   Sistrurus miliarius (pigmy rattlesnake)
SOURCE     Sistrurus miliarius
ORGANISM   Sistrurus miliarius
REFERENCE   Sheppard, P.O. and Bishop, P.D.
AUTHORS     Rattlesnake venom gland proteins
TITLE       Patent: WO 0214364-A 4 21-FEB-2002;
JOURNAL     Zymogenetics, Inc. (US)
FEATURES   Location/Qualifiers
            1..721
            /organism="Sistrurus miliarius"
            /mol_type="genomic DNA"
            /db_xref="taxon:8758"
            91..549
            /note="Zsnnk3"
            CDS

/product="agglucetin-alpha 1 subunit"
BASE COUNT      166 a   163 c   178 g   159 t
ORIGIN
Query Match      57.4%; Score 312; DB 5; Length 666;
Best Local Similarity 78.7%; Pred. No. 4e-82;
Matches 428; Conservative 0; Mismatches 100; Indels 16; Gaps 4;
QY 9 TTGCTACAAAGCTCTTCAAAACAACTTAAGACCTGGACAGATGCAGAGCTTCTGCACGAA 68
Db 128 TTGCTACCGGCTCTCAAACTCTCAAGCTCTGGATGATGCAGAGAGTTCTGCACGGA 187
QY 69 GCAGGTGAACCGGGGGCATCTGGTCTCTATCGAAAGTCCGGAGAGCAGACTTTTGTGGG 128
Db 188 GCGGCCGAAAGCGGGCATCTGGTCTCTATCGAAAGCGGGAGAGAGACTTTTGTGGC 247
QY 129 CCAGTTGATTGCTCAGAAGATAAGTCAAGCCAAATCATCTGATCGGACTGAGGSC 188
Db 248 CCAGCTGGTCTCTGAGAAACAAACAGACGACA-----ATGTCCTGGCTTGGACTGAAGAT 301
QY 189 TCAAACAAAGAAAGCAATGAGCATAGAGTGGAGCGATGGCTCCAGCATCAGTTATGA 248
Db 302 TCAAGCAAGAGCAGCAATGCAGACAGAGTGGACCGATGGCTCCAGGTCAGTTATGA 361
QY 249 GAATGGATTGAAGAGAAATCCAAAAGTGCTTGGGGTGCAATAGAGACAGGGTTTCA 308
Db 362 GAATTTTCTGAATACCAATCCAAAAGTGTTTGTGTGGAAGAAACACAGGGTTTGC 421
QY 309 TAAGTGGAGATTTTACTGTGAACACAGATCCTTTTGTCTGCA-----GGCATA 362
Db 422 CACATGGCTCAATCTTAATTTGGGAAGTGAATATGCTTTCGTCGAGATCCCGCCATA 481
QY 363 GTCTGAAGATCAGCTGATTTGAAGTCTGGAGAGCAAGGAAGCCCGCCACCCATCCCCC 422
Db 482 GTGTTAAGATCAGCTGAGTGAAGTCTGGAGAGCAAGGAAGCCCGCCACCCCA--GCC 539
QY 423 AACCTGCTAGCCACAATCTCTGCTATGCATGCAACCTTTTGTCTAAGGATGCTCTGTAGC 482
Db 540 ACCCCCTGCTGCGGCTATCTGCTCTGCAACCTTTTGTCTCAACGATGCTCTCTGTAGC 599
QY 483 TGGATCTGTTGTTGCTGCTCTGATGGCGCGGAG--TCAATAAATCTGCTAGCTGCTGA 540
Db 600 TGGATCTGTTTGTGCTCTGCTGATGGCGCGGAGTCCAAATAAATCTGCTAGCAAAA 659
QY 541 AAAA 544
Db 660 AAAA 663

RESULT 11
AF102901
LOCUS      AF102901
DEFINITION Deinagkistrodon acutus agkicetin alpha subunit precursor (GP1bA1)
ACCESSION AF102901
VERSION    AF102901.1 GI:11967284
KEYWORDS   mRNA, complete cds.
SOURCE     Deinagkistrodon acutus
ORGANISM   Deinagkistrodon acutus
REFERENCE   Lepidosauria; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE       Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
JOURNAL     Viperidae; Crotalinae; Deinagkistrodon.
            1 (bases 1 to 704)
            Chen, Y.L., Hong, T.M., Chang, T. and Tsai, I.H.
            cDNA sequence and functional characterization of glycoprotein
            1b-binding protein from the venom of Deinagkistrodon acutus
            Unpublished
            2 (bases 1 to 704)
            Chen, Y.L. and Tsai, I.H.
            Direct Submission
            Submitted (30-OCT-1998) Biological Chemistry, Academia Sinica,
            Taipei, Taiwan, R.O.C.
            Location/Qualifiers
            1..704
            /organism="Deinagkistrodon acutus"
            /mol_type="mRNA"
            source
```

```

/db xref="taxon:36307"
<1..704
/gene="GP1B1"
CDS
61..528
/gene="GP1B1"
/note="antagonist of vWF-induced platelet agglutination"
/codon_start=1
/product="agkicetin alpha subunit precursor"
/protein_id="AAG42040.1"
/db xref="GI:11967285"
/translation="MRFLFVSGLIVFLSLSGTAADCLPMSSYIRFCYQPKLIK
TWDAERFCTEQAGHGLVFSAREADVAGLVSENIKIPYVWIGLRVQNEGQCS
SKWSDSKSVSYENLVFPFSKKFVLKDDTGFRTWENVYGLKHVFNCKYLKPR"
61..129
sig_peptide
130..525
mat_peptide
/gene="GP1B1"
/product="agkicetin alpha subunit"
175 a 167 c 187 g 175 t
BASE COUNT
ORIGIN
Query Match 56.2%; Score 306; DB 5; Length 704;
Best Local Similarity 77.7%; Pred. No. 2.5e-80;
Matches 426; Conservative 0; Mismatches 105; Indels 17; Gaps 4;
QY 5 GGCAATGCTACAGGTCTTTCAACAATCTAAGACCTGGACAGATGCAGAGAGCTTCTGCA 64
Db |||||
161 GGTGTTGCTACAGCCCTTCAAGCTCTCAAGACCTGGAGATGCAGAGAGCTTCTGCA 220
QY 65 CGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAGCAGACTTTG 124
Db |||||
221 CGAGCAGCGAATGGGGGGCATCTGGTCTCTTTCGAAAGCGCAGAGAGCAGACTTTG 280
QY 125 TGGCGCAGTGTGATGCTCAGAAGATAAAGTCAGCCAAATCCATGCTCGATCGAGTGA 184
Db |||||
281 TGGCGGGTGTCTCTGAGAACATAAA---AATCAACCCCTATGCTGGATTGGACTGA 337
QY 185 GGGCTCAAAACAGAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATCAGTT 244
Db |||||
338 GGGTTCAAAACAGAGCAGCAATGCAGCTCTAAGTGGAGTATAGTCCAAAGTCAGTT 397
QY 245 ATGAGAAATGGAATGAAGAGAAATCAAAAGTGTCTGGGGTGCAATAGAGCAGGCT 304
Db |||||
398 ATGAGAACTGGTTGAACATTTTCCAAAAGTGTCTTGTGCTGAAAAGGACACAGGCT 457
QY 305 TTCTAAGTGGAGAAATTTTACTGTGAACAACAGATCCTTTTCTCTCGA-----GG 358
Db |||||
458 TTGCTACGTGGAGAAATGTTACTGTGACATAAACATGTTTTCATGTGCAAGTACCTGA 517
QY 359 CATAGTCTGAAGATCCAGCTGATGGAAGTCTGGAAGCAAGAAAGCCCCCACCACATC 418
Db |||||
518 AACCGGTAAAGATCGGCTGCGAAGTCTGGAAGCAAGAAAGCCCCCACCACAC 577
QY 419 CCCCAACCTCCCTAGCCACATCTCTGCTATGCAACCTTTGCTCAACGATGCTCTG 478
Db |||||
578 CCCCAGCC-----TGCCGCTATGCTTCTCTGCAACCTTTGCTCAACGATGCTCTG 631
QY 479 TAGCTGATCTGGTGTCTGCTGCTGATGGCCGGAAG--TCAATAAATCTGCTAGC 536
Db |||||
632 TAGCTGATCTGGTGTCTGCTGCTGATGGCCGGAAGTCCAAATAAATCTGCTAGC 691
QY 537 CTGAAAAA 544
Db |||||
692 AAAAAAAA 699

```

```

RESULT 12
AV149341
LOCUS Trimeresurus flavoviridis flavocetin-A alpha chain mRNA linear VRT 24-NOV-2002
DEFINITION Trimeresurus flavoviridis flavocetin-A alpha chain mRNA, complete
cda.
ACCESSION AV149341
VERSION AV149341.1 GI:25245560

```

KEYWORDS

SOURCE
ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1..681

/organism="Trimeresurus flavoviridis"

/mol_type="mRNA"

/db_xref="taxon:88087"

/tissue_type="venom glands"

37..513

/note="partial peptide sequencing"

/codon_start=1

/evidence="experimental"

/product="flavocetin-A alpha chain"

/protein_id="AAN72438.1"

/db_xref="GI:25245561"

/translation="MERLIFVSGLLVILSLSGTADFDICPGMSAYDRYCYQAFSK

PKRWDAFSCFEEGVKTSKLVSISSGDFVAOLVAEKIKTSFOVWIGLRONKEQ

QCRSWSVSSSVNENLVKFKSKYALKKGLTFRWNVYGTENPFVCKTPEC"

BASE COUNT 170 a 164 c 182 g 165 t

ORIGIN

Query Match 54.7%; Score 297.4; DB 5; Length 681;

Best Local Similarity 76.7%; Pred. No. 9.4e-78;

Matches 418; Conservative 0; Mismatches 116; Indels 11; Gaps 4;

```

QY 1 GAAGGGCATTCGTACAGGTCTTCAACAATCTAAGACCTGGAGATGCAGAGAGCTTTC 60
Db |||||
139 GATCGGTATGCTACAGGCTTTCAGTAAACCGAAAACTGGGAAGATGCAGAGATTC 198
QY 61 TGCAGAGACAGGTCAACGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAAGCAGAC 120
Db |||||
199 TCGGAGAGGGGTGAAGACCTCGCATTTGGTCTCAATCGAAAGCTCCGGAAGAGAC 258
QY 121 TTTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAATCCATGCTCTGGATCGA 180
Db |||||
259 TTCGTGGCCAGCTGGTGGTGAAGATAAAGACTTCTTTTTCAGTATGCTGGATTGA 318
QY 181 CTGAGGGCTCAAAACAGAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGATC 240
Db |||||
319 CTGAGATTCAAAACAGAAAGCAATGCAGGTGGAGTGGAGCGATGCTCCAGTGTCT 378
QY 241 AGTTATGAGAAATGGATTGAAGAAATCCAAAAAGTCTTTGGGGTGCACATAGAGACA 300
Db |||||
379 AATTATGAGAACTGGTTAAACAATTTTCCAAAAATGTTATGCTGAAAAAAGGGACA 438
QY 301 GGGTTTCATAAGTGGGAGAAATTTTATCTGTGAAACAAAGATCCTTTTGTCTGGCA --- 356
Db |||||
439 GAGCTTCGACGCTGGTTCAACGTATCTGTGTGAACAGAAAAATCTTTTGTCTGCAAGTAC 498
QY 357 --GGCATAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAGCAAGAGAGCCCCCACC 414
Db |||||
499 ACGCCAGATGTTAAGATCCAGCTGAGTGAAGTCTGGAGAGCAAGAGAGCCCCCAGCC 558
QY 415 CATCCCCAACCCCTGCTAGCACAATCTCTGTCTATGCACCCCT--TTGCTCAACGATGCT 473
Db |||||
559 CA--CCCCACCCCTGCTGCGCAATCTCTGCTCTGCCCCCTCTCATTCATGATGATGCT 616

```

Trimeresurus flavoviridis

Trimeresurus flavoviridis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;

Viperidae; Crotalinae; Trimeresurus.

1 (bases 1 to 681)

Shin,Y., Okuyama,I., Hasegawa,J. and Morita,T.

Molecular cloning of glycoprotein Ib-binding protein, flavocetin-A,

which inhibits platelet aggregation

Thromb. Res. 99 (3), 239-247 (2000)

20402624

10942790

2 (bases 1 to 681)

Shin,Y., Okuyama,I., Hasegawa,J. and Morita,T.

Direct Submission

Submitted (05-SEP-2002) Department of Biochemistry, Meiji

Pharmaceutical University, 2-522-1, Noshio, Kiyose, Tokyo 204-8588,

Japan

Location/Qualifiers

1..681

/organism="Trimeresurus flavoviridis"

/mol_type="mRNA"

/db_xref="taxon:88087"

/tissue_type="venom glands"

37..513

/note="partial peptide sequencing"

/codon_start=1

/evidence="experimental"

/product="flavocetin-A alpha chain"

/protein_id="AAN72438.1"

/db_xref="GI:25245561"

/translation="MERLIFVSGLLVILSLSGTADFDICPGMSAYDRYCYQAFSK

PKRWDAFSCFEEGVKTSKLVSISSGDFVAOLVAEKIKTSFOVWIGLRONKEQ

QCRSWSVSSSVNENLVKFKSKYALKKGLTFRWNVYGTENPFVCKTPEC"

BASE COUNT 170 a 164 c 182 g 165 t

ORIGIN

Query Match 54.7%; Score 297.4; DB 5; Length 681;

Best Local Similarity 76.7%; Pred. No. 9.4e-78;

Matches 418; Conservative 0; Mismatches 116; Indels 11; Gaps 4;

```

QY 474 CTCTGTAGTGGATCTGGTGTCTGCTCTGATGGCCGGAAG--TCAATAAATTCG 531
Db 617 CTCTGTAGTGGATCTGGTGTCTGCTCTGATGGCCGGAAGTCCATAAATTCG 676

QY 532 CTAGC 536
Db 677 CTAGC 681

RESULT 13
CDRNACVXA 800 bp mRNA linear VRT 30-JUL-1998
LOCUS Crotalus durissus mRNA for convulxin alpha.
DEFINITION Y16348
ACCESSION Y16348.1 GI:3367645
VERSION convulxin alpha; CVX alpha gene.
KEYWORDS Crotalus durissus (tropical rattlesnake)
SOURCE Crotalus durissus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
Viperidae; Crotalinae; Crotalus.
REFERENCE 1
AUTHORS Leduc,M. and Bon,C.
TITLE Cloning of subunits of convulxin, a collagen-like
platelet-aggregating protein from Crotalus durissus
terrificus
venom
JOURNAL Biochem. J. 333 (Pt 2), 389-393 (1998)
MEDLINE 98324901
PubMed 9657980
REFERENCE 2 (bases 1 to 800)
Direct Submission
AUTHORS Leduc,M.
TITLE Submitted (27-JAN-1998) M. Leduc, Institut Pasteur, Unite des
Venins, 25 Rue du Dr Roux, 75724 Paris cedex 15, Paris, FRANCE
JOURNAL
FEATURES
source
1..800
/organism="Crotalus durissus"
/mol_type="mRNA"
/sub_species="terrificus"
/db_xref="taxon:8731"
/clone="pt3r7"
/tissue_type="venom gland"
1..800
/gene="CVX alpha"
1..800
/gene="CVX alpha"
1..800
/product="convulxin alpha"
133..609
/gene="CVX alpha"
/evidence=experimental
/codon_start=1
/product="convulxin alpha"
/protein_id="CAA76181.1"
/db_xref="GI:3367646"
/db_xref="SPTREMBL:O93426"
/translation="MGRIFFVSFGLLVFLSLSGTGAGLHCPSDWYYDDHOCYRIFNE
EMWDEAWFCTKQAKGHLVSIKAKEADFAVMVTONIESFHSVIGLKVQNKKEK
QCSTKWSGSSVSVDNLDLYTKCSLLKKEGFKWKFVASCIGIKIPFVKFPFQC"
211..606
/mat_peptide
/gene="CVX alpha"
/product="convulxin alpha"
/evidence=experimental
183 c 218 g 196 t

BASE COUNT 203 a 183 c 218 g
ORIGIN

Query Match 53.0%; Score 288.2; DB 5; Length 800;
Best Local Similarity 75.2%; Pred. No. 5.4e-75;
Matches 416; Conservative 0; Mismatches 123; Indels 14; Gaps 4;

QY 1 GAAGGCATTTGCTACAGGCTCTTCAACATCTTAAGACCTGGACAGATGCAGAGACTTC 60
Db 235 GATCAGCATTTGCTACCGGATCTTCAATGAAGAGATGAACCTGGGAAGATGCAGAGTGGTTC 294

```

```

QY 61 TGCACGAGCAGGTGAACGGGGGCGATCTGGTCTCTATCGAAAGCTCCGAGAGCAGAC 120
Db 295 TGCACGAAGCAGCGAAGCGCGCGATCTCGTCTCTATCAAAAGCGCAAGAGCAGAC 354

QY 121 TTTGTGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAATCCATGCTCTGGATCGGA 180
Db 355 TTTGTGGCCCTGGATGGTCACTCAGAACATAGAGGATCCTTTTCCCATGCTCGATTGGA 414

QY 181 CTGAGGCTCAAAACAAAGAAACCAATGCAGCATAGATGAGTGGAGCGATGGCTCCAGATC 240
Db 415 CTGAGGCTTCAAAACAAAGAAACCAATGCAGCACGAAGTGGAGCGATGGCTCCAGCGTC 474

QY 241 AGTTATGAGAAATGGATTTGAAGAAGATCAAAAAGTGTCTTGGGGTGCACATAGAGACA 300
Db 475 AGTTATGACAACTTGTGGATCTATATATTACAAAGTGTAGTCTGCTGAAAAAAGAGACA 534

QY 301 GGGTTTCATAAGTGGGAGAAATTTTACTGTGAACAACAAGATCCTTTTGTCTGCGA---- 356
Db 535 GGGTTTCGTAAGTGGTTCGTTGCTAGCTGTATAGGAAGATTCTTTTCTGCTGCAAGTTC 594

QY 357 --GGCATAGTCTGAAGATCCAGTGTGATTTGAAGTCTGGAGAAGCAAGAGCCCCCACC 414
Db 595 CCGCCACAGTGTAAAGTACCAGCTTTGTGAAGTCTGGAGAAGCAAGAGCCCCCACC-- 652

QY 415 CATCCCAACCCCTGCTAGCCAAATCTGCTATGCACCT--TTGCTCAACGATGCT 473
Db 653 ---TCCCACCCCGCCCTGCCCAATCTCTGCTCTGACCTGTCTTCCATGGTGTCT 709

QY 474 CTCTGTAGTGGATCTGGTGTGCTGCTCTGATGGCGCGAAG--TCAATAAATTCG 531
Db 710 CTCTGTGGTGATCTGGTGTGCTGCTCTGATGGCGCGAAGTCCAAATAATTCG 769

QY 532 CTAGCCTGAAAAA 544
Db 770 CTAGCATGAAAAA 782

RESULT 14
LOCUS AX427210 725 bp DNA linear PAT 18-JUN-2002
DEFINITION Sequence 10 from Patent WO0214364.
ACCESSION AX427210
VERSION AX427210.1 GI:21530572
KEYWORDS Sistrurus miliarius (pigmy rattlesnake)
SOURCE Sistrurus miliarius
ORGANISM Sistrurus miliarius
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
Viperidae; Crotalinae; Sistrurus.
REFERENCE 1
AUTHORS Sheppard,P.O. and Bishop,P.D.
TITLE Rattlesnake venom gland proteins
JOURNAL Patent: WO 0214364-A 10 21-FEB-2002;
ZymoGenetics, Inc. (US)
FEATURES
source
1..725
/organism="Sistrurus miliarius"
/mol_type="genomic DNA"
/db_xref="taxon:8758"
88..564
/note="Zenk5"
/codon_start=1
/protein_id="CAD35501.1"
/db_xref="GI:21530573"
/translation="MGRIFFVSFGLLVFLSLSGTGADFNCPGFWFAYDQCYRVIKR
LXTWDDARFCEQAKGHLASVENDEAVFLQIVAAANIKONQYVYVGLRIONKQ
QCSTKWSGSSVSVDNLDLYTKCSLLKKEGFKWKFVASCIGIKIPFVKFPFPEC"
179 a 171 c 200 g 175 t

BASE COUNT 179 a 171 c 200 g
ORIGIN

Query Match 52.9%; Score 288; DB 6; Length 725;
Best Local Similarity 75.9%; Pred. No. 6.2e-75;

```

Matches 416; Conservative 0; Mismatches 110; Indels 22; Gaps 4;
QY 1 GAAGGCGATGCTACAGGCTCTTCAACAATCTAAGACCTGGACAGATGAGAGAGCTTC 60
Db 190 GATCAGTATGCTACAGGCTCATCAACGACTCAAGACCTGGACAGATGAGAGAGCTTC 249
QY 61 TGCACCAAGCAGGTGAACGGGGGCGCATCTGGCTCTCTATCGAAAGCTCCGGAGAGCAGAC 120
Db 250 TGCTCGGAGCAGCGAAGCGGGCGCATCTGGCTCTCTATCGAAAGCTCCGGAGAGCAGAC 309
QY 121 TTGTGGGCGCAGTTGCTCAGAGATTAAGTCAAGCAATCTAAGACCTGGACAGATGAGAGAGCTTC 180
Db 310 TTGTGGGCGCAGTTGCTCAGAGATTAAGTCAAGCAATCTAAGACCTGGACAGATGAGAGAGCTTC 369
QY 181 CTGAGGCTCAAAACAAAGAAAGCAATGAGATGAGAGAGCTCCAGCATC 240
Db 370 CTGAGGATTCAAAACAAAGGACAGCAATGAGATGAGAGAGCTCCAGCATC 429
QY 241 AGTTATGAGAAATGGAATGAAGAGAAATCCAAAAGTCTTGGGGTGACATAGAGACA 300
Db 430 AGTTATGAGAAATGGAATGAAGAGAAATCCAAAAGTCTTGGGGTGACATAGAGACA 489
QY 301 GGGTTTCATAGTGGGAGAAATTTTACTGTGAACAAGATCCCTTTTCTGCGA---- 356
Db 490 GAGTTCTTCAATGTTAATCAATGCTGACTCGGAGAAAGAAACCTTTTCTGCGAAGTTTC 549
QY 357 --GGCATAGTCTGAAGATCCAGCTGATGAGTCTGGAGAAAGCAAGAGAGAGAGAGAG 414
Db 550 CGGCGAGAGTGTAAAGTCCGCTGTGAGTCTGGAGAGCAAGAGATCCCGCCAC 609
QY 415 CATCCCCCAACCTGCTAGCAGCAATCTCTGCTGCACTCTGCACTCTGCACTGCACTGCT 473
Db 610 CGGCC-----TGCACAACTCTCTGCTGCACTCTGCACTCTGCACTGCACTGCT 656
QY 474 CTCTGTAGTGTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 531
Db 657 CTCTGTGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 716
QY 532 CTAGCCTG 539
Db 717 CTAGCATG 724

RESULT 15
AX427207
LOCUS AX427207 580 bp DNA linear PAT 18-JUN-2002
DEFINITION Sequence 7 from Patent WO0214364.
ACCESSION AX427207
VERSION AX427207.1 GI:21530569
KEYWORDS
SOURCE
ORGANISM
Sistrurus miliarius (pigmy rattlesnake)
Sistrurus miliarius
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Lepidodactylia; Squamata; Scleroglossa; Serpentes; Colubroidae;
Viperidae; Crotalinae; Sistrurus.
1
Sheppard, P.O. and Bishop, P.D.
Rattlesnake venom gland proteins
Patent: WO 0214364-A 7 21-FEB-2002;
ZymoGenetics, Inc. (US)
FEATURES
Location/Qualifiers
1..580
/organism="Sistrurus miliarius"
/mol_type="genomic DNA"
/db_xref="taxon:8758"
<3..437
/note="Zenk4"
/codon_start=1
/protein_id="CAD35500.1"
/db_xref="GI:21530570"
/translation="IRNEGSGTADFCPSDWAYDQCYRVIKOLRTWDDAERFCSEQ
AKGSHLVSIKSDGAEAFVQALVAENIKYDVWIGLRIQEEKQCSKRWSDSSVNY
ENLIKATKKCFGLKKETGFTWRNVHCTQONLFMCKFPPEC"
CDS

BASE COUNT 154 a 133 c 163 g 130 t
ORIGIN
Query Match 52.5%; Score 285.8; DB 6; Length 580;
Best Local Similarity 75.2%; Pred. No. 2.9e-74;
Matches 413; Conservative 0; Mismatches 97; Indels 39; Gaps 3;
QY 1 GAAGGCGATGCTACAGGCTCTTCAACAATCTAAGACCTGGACAGATGAGAGAGCTTC 60
Db 63 GATCAGTATGCTACAGGCTCATCAACGACTCAAGACCTGGACAGATGAGAGAGCTTC 122
QY 61 TGCACCAAGCAGGTGAACGGGGGCGCATCTGGCTCTCTATCGAAAGCTCCGGAGAGCAGAC 120
Db 123 TGCTCGGAGCAGCGAAGCGGGCGCATCTGGCTCTCTATCGAAAGCTCCGGAGAGCAGAC 182
QY 121 TTGTGGGCGCAGTTGATGCTCAGAGATTAAGTCAAGCAATCTAAGACCTGGACAGATGAGAGAGCTTC 180
Db 183 TTGTGGGCGCAGTTGATGCTCAGAGATTAAGTCAAGCAATCTAAGACCTGGACAGATGAGAGAGCTTC 242
QY 181 CTGAGGCTCAAAACAAAGAAAGCAATGAGATGAGAGAGCTCCAGCATC 240
Db 243 CTGAGGATTCAAAACAAAGGACAGCAATGAGATGAGAGAGCTCCAGCATC 302
QY 241 AGTTATGAGAAATGGAATGAAGAGAAATTTTACTGTGAACAAGATCCCTTTTCTGCGA---- 356
Db 303 AATTATGAGAACTGATTAACATCGGACCAAAAGTGTGTTGGGCTGAAAAGAGAGACA 362
QY 301 GGGTTTCATAGTGGGAGAAATTTTACTGTGAACAAGATCCCTTTTCTGCGA---- 422
Db 363 GGGTTTCGACGCTGGCGCAATGTTCACTGTACACAACAATACTTTTCTGCAAGTTC 422
QY 357 --GGCATAGTCTGAAGATCCAGCTGATGAGTCTGGAGAGCAAGAGAGAGAGAGAGAG 414
Db 423 CGGCCAGAGTGTAAAGTCCGCTGTGAGTCTGGAGAGCAAGAGAGAGAGAGAGAGAG 482
QY 415 CATCCCCCAACCTGCTAGCAGCAATCTCTGCTATGACACCTTTGCTCAACGAGATGCTC 474
Db 483 CACCGC-----CACCTTTGCTCAACGAGATGCTC 511
QY 475 TCTGTAGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 532
Db 512 TCTGTAGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 571
QY 533 TAGCCTGA 541
Db 572 TAGCCTGA 580

Search completed: December 8, 2003, 12:01:59
Job time : 3183 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2003, 09:08:47 ; Search time 321 Seconds
(without alignments)
4574.749 Million cell updates/sec

Title: US-09-938-114-1
Perfect score: 544
Sequence: 1 GAAGGGCATTGCTCAAGGT.....ATTCTGCTAGCTGAAAAA 544

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 19Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	544	100.0	544	25	ABX93668 cDNA encoding Dein
2	306.2	56.3	721	24	AAD32055 Pigmy rattlesnake
3	288	52.9	725	24	AAD32059 Pigmy rattlesnake
4	285.8	52.5	580	24	AAD32057 Pigmy rattlesnake
5	276.2	50.8	632	24	AAL42015 Korean adder snake
6	266	48.9	601	22	AAI171876 Snake venom blood
7	226.4	41.6	690	16	AAQ89309 Snake venom antith
8	226.4	41.6	690	21	AAC61144 DNA encoding a sna

9	226.4	41.6	690	24	AAK99834	DNA encoding the a
10	186.6	34.3	456	24	AAD32056	Pigmy rattlesnake
11	175	32.2	432	24	AAD32058	Pigmy rattlesnake
12	166	30.5	474	24	AAD32060	Pigmy rattlesnake
13	144	26.5	633	22	AAI171877	Snake venom blood
14	102.6	18.9	272	16	AAQ89308	Snake venom antith
15	102.6	18.9	272	21	AAC61143	Rattlesnake venom
16	98.8	18.2	583	24	AAL42016	Korean adder snake
17	82.4	15.1	483	24	AAD32053	Pigmy rattlesnake
18	81.2	14.9	716	17	AAT64829	Snake venom protei
19	64.2	11.8	453	24	AAD32054	Pigmy rattlesnake
20	58	10.7	454	25	ABX93674	DNA encoding Dein
21	39	7.2	10598	23	ABL27241	Drosophila melanog
22	39	7.2	10629	22	AAH28279	Nucleotide sequenc
23	39	7.2	15667	23	ABL27240	Drosophila melanog
24	36.8	6.8	1597	21	AAC59784	Human secreted pro
25	36.8	6.8	1597	22	AAF24170	Human secreted pro
26	36.8	6.8	128600	24	ABK83461	Human cDNA differe
27	35.4	6.7	498	24	ABK46204	cDNA encoding colo
28	36	6.6	492	22	AAI64154	Human Zlec1 encodi
29	36	6.6	492	24	AAD27975	Human lectin Zlec1
30	34.8	6.4	461	24	ABK46210	cDNA encoding colo
31	34.8	6.4	509	24	ABV97585	Human pancreatic c
32	34.6	6.4	2949	21	AAD00024	Degenerate DNA enc
33	34.6	6.4	2949	24	AAD37455	Human Zfsta2 degen
34	34.4	6.3	1221	23	AAH67247	DNA encoding novel
35	34.4	6.3	1222	23	AAH66042	DNA encoding novel
36	34.4	6.3	2581	23	AAH64652	DNA encoding novel
37	34.2	6.3	441	11	AAQ05622	Sequence encoding
38	34.2	6.3	493	24	ABK46235	cDNA encoding colo
39	34.2	6.3	495	10	AAH91109	Human reg cDNA. H
40	34.2	6.3	498	9	AAH81962	Sequence of human
41	34.2	6.3	800	22	AAH75077	Human colon associ
42	34.2	6.3	843	21	AAH75090	Human cancer associ
43	34.2	6.3	843	22	AAH75090	Human colon associ
44	34	6.2	790	11	AAQ05300	Bovine pancreatic
45	33.8	6.2	585	24	ABK46227	cDNA encoding colo

ALIGNMENTS

RESULT 1
ABX93668
ID ABX93668 standard; cDNA; 544 BP.

XX

AC ABX93668;

XX 02-JUN-2003 (first entry)

XX cDNA encoding Deinagkistrodon acutus antithrombosis enzyme alpha chain.
XX Antithrombosis; ss; PCR; alpha chain; fibrin hydrolysis; blood clot;
XX platelet aggregation; vaso-occlusive disorder; thromboembolic disorder;
XX myocardial infarction; restenosis; cancer; neurodegenerative disease;
XX angiopathic thrombosis; cerebral thrombosis; thromboangiitis obliterans;
XX ischaemic cerebral vascular disease; unstable angina; acute thrombosis;
XX unstable stenocardia; pulmonary embolism; deep vein thrombosis; oedema;
XX peripheral arterial occlusion; stroke; atherosclerosis; inflammation;
XX thrombosis.

OS Deinagkistrodon acutus.

XX Key Location/Qualifiers
XX CDS 1..363
XX /tag= a
XX /product= "Antithrombosis enzyme alpha chain"
XX /partial
XX /transl_except= (pos:244..246,aa:Lys)
XX /note= "No start codon given"

US6489451-B1.

CC weight of expressed polypeptides as a ratio to total protein expressed,
 CC identifying peptide cleavage sites, coupling amino and carboxyl terminal
 CC tags, mass spectrometry, circular dichroism to determine conformation or
 CC affinity chromatography columns to purify the protein, cloning or
 CC sequencing. The present sequence is Sistrurus miliarius venom gland
 CC protein, Zsnk3 gene.
 XX SQ Sequence 721 BP; 192 A; 167 C; 203 G; 159 T; 0 other;
 Query Match 56.3%; Score 306.2; DB 24; Length 721;
 Best Local Similarity 78.8%; Pred. No. 5e-86; Indels 0; Gaps 0;
 Matches 365; Conservative 0; Mismatches 98; Indels 0; Gaps 0;
 QY 1 GAAGGGCTTCTCAAGGCTTCTCAACAACTTAAGACCTGACAGATCGACAGAGCTTC 60
 DB 187 GATCAGCATTCACAGGCTTCTCAACAACTTAAGACCTGACAGATCGACAGAGCTTC 246
 QY 61 TGCAGGAAGCGGTGAACGGGGGGCATCTGCTCTATTCGAAAGTCCGGGAAGCAGAC 120
 DB 247 TGCTCGAGCAGCGCGGGGGCATCTGCTCTATTCGAAAGTCCGGGAAGCAGAC 306
 QY 121 TTTGTGGCCAGTTGATGCTCAGAAGATAAAGTCAGCCAAAATCCATGCTGGATCGGA 180
 DB 307 TTTGTGGCCAGTGGTCCCTGAGAACAGGAGGAGCCATCTCTATTCGATCGGA 366
 QY 181 CTGAGGGCTCAAAAACAAAGAAAGCAATCGACATAGAGTGGAGCGATGGCTCCAGCATC 240
 DB 367 CTGAGGGTTCAGGCAAGAGAGCAATCGACGCGGAAGTGGAGCGATGGCTCCAGCGTC 426
 QY 241 AGTTATGAGATTGATTAAGAGAAATCCAAAAGTGTCTGGGGTGCACATAGAGACA 300
 DB 427 AGTTATGAGATTGATTAAGAGCAATCCAAAATGCTTTGGGCTGCACAAAGGCACA 486
 QY 301 GGGTTTCATAGTGGGAGAAATTTTACTGTGAACAAAGATCCCTTTGCTCGCAGGCA 360
 DB 487 AATTATCATAGTGGTCAATATTCTGTGAGAAATAATCTTTGCTCGCAGGCA 546
 QY 361 TAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAGCAAGGAAGCCGCCCATGCC 420
 DB 547 TAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAGCAAGGAAGCCGCCCATGCC 606
 QY 421 CCAACCCCTGCTAGCCACATCTGCTATGACCCCTTGCTC 463
 DB 607 CCCCACCTGCTGATCTGTAGTGGGATCTGTTCTGCTGCTC 649
 RESULT 3
 AAD32059
 ID AAD32059 standard; DNA; 725 BP.
 XX AC AAD32059;
 XX DT 18-JUN-2002 (first entry)
 XX DE Pigmy rattlesnake venom gland protein, Zsnk5 gene.
 XX KW Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;
 XX KW platelet aggregation; gene; Zsnk5; ds.
 XX OS Sistrurus miliarius.
 XX FH Key Location/Qualifiers
 XX CDS 88..564
 XX FT /*tag= a
 XX FT /product= "Zsnk5 protein"
 XX FT sig_peptide 88..156
 XX FT /*tag= b
 XX FT mat_peptide 157..561
 XX FT /*tag= c
 XX FT /product= "Mature Zsnk5 protein"
 XX PN WO200214364-A2.
 XX XX

PD 21-FEB-2002.
 XX 13-AUG-2001; 2001WO-US25310.
 XX 14-AUG-2000; 2000US-225072P.
 PR 14-AUG-2000; 2000US-225087P.
 PR 15-AUG-2000; 2000US-225489P.
 PR 15-AUG-2000; 2000US-225490P.
 PR 20-DEC-2000; 2000US-356997P.
 XX (ZYMO) ZYMOGENETICS INC.
 PA Sheppard PO, Bishop PD;
 PI WPI; 2002-269180/31.
 XX P-PSDB; AAE20181.
 DR New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which
 PT affect blood coagulation and platelet aggregation system, useful in
 PT therapy and diagnostics, or as tools in the study of genetics or
 PT molecular biology
 PS Claim 5; Page 77; 79pp; English.
 XX The invention relates to new pigmy rattlesnake (Sistrurus miliarius)
 CC venom gland proteins, which affect blood coagulation and platelet
 CC aggregation system. The polypeptides, which affect blood coagulation and
 CC platelet aggregation system, are useful in therapy and diagnostics. The
 CC polypeptides are also useful as an educational tool in laboratory
 CC practical kits for courses related to genetics and molecular biology,
 CC protein chemistry and antibody production and analysis. The
 CC polynucleotide or polypeptide can be used as standards or as unknowns
 CC for testing purposes. The polypeptides are also useful in identifying
 CC proteins by western blotting, protein purification, determining the
 CC weight of expressed polypeptides as a ratio to total protein expressed,
 CC identifying peptide cleavage sites, coupling amino and carboxyl terminal
 CC tags, mass spectrometry, circular dichroism to determine conformation or
 CC affinity chromatography columns to purify the protein, cloning or
 CC sequencing. The present sequence is Sistrurus miliarius venom gland
 CC protein, Zsnk5 gene.
 XX SQ Sequence 725 BP; 179 A; 171 C; 200 G; 175 T; 0 other;
 Query Match 52.9%; Score 288; DB 24; Length 725;
 Best Local Similarity 75.9%; Pred. No. 2.8e-80; Indels 22; Gaps 4;
 Matches 416; Conservative 0; Mismatches 110;
 QY 1 GAAGGGCTTCTCAAGGCTTCTCAACAACTTAAGACCTGACAGATCGACAGAGCTTC 60
 DB 190 GATCAGTATTGCTACAGGTCATCAACGACTCAAGACCTGGGACGATCGACAGCGTTC 249
 QY 61 TGCAGGAAGCGGTGAACGGGGGGCATCTGCTCTATTCGAAAGTCCGGAGAGCAGAC 120
 DB 250 TGCTCGAGCAGCGGAAGGGCGGCATCTGGGCTCTGTGAAAACGATGAAGAGCGTC 309
 QY 121 TTTGTGGCCAGTTCATTCCTCAGAGATAAAGTCAAGCAATCCATCTCTGATCGCA 180
 DB 310 TTTCTGGCCAGTTCGCTCGCAACATAAGCAAAACCAATCTATCTCTGATTTGA 369
 QY 181 CTGAGGGCTCAAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
 DB 370 CTGAGGATTCAAAAACAAAGAGCAGCAATGCAGCAGAGTGGAGCGATGGCTCCAGCGTC 429
 QY 241 AGTTATGAGATTGATTAAGAGAAATCCAAAAGTGTCTTGGGGTGCACATAGAGACA 300
 DB 430 AGTTATGAGAACTGGTTAAATCAATTCAAAAGTGTCTTGGGCTGAAAAGAGACA 489
 QY 301 GGGTTTCATAGTGGGAGAAATTTTACTGTGAACAAAGATCCCTTTCTCTGCGA --- 356
 DB 490 GAGTTTCTTCAATGGGTACAACTGACTGCGAAGAAAAAACCTTTTCTGCAAGTTC 549
 QY 357 --GGCATAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAGCAAGAGAGCCGCCACCC 414

Db 550 CCGCAGAGTGTTAAGATCCGGTGTGTAAGTCTGGAGAAGCAAGAAATCCCCCCCCC 609
 QY 415 CATCCCCCAACCTGCTAGCCCAATCTCTGCTATGACACCT-TTCCTCAACGGATGCT 473
 Db 610 CCGCC-----TCCCAATCTCTGCTCTGCACTCTCTCAATTCATGGATGCT 656
 QY 474 CTCTGAGTGTGATCTGTTGCTGCTCTGCTGATGGCCCGAAG--TCAATAAATCTGC 531
 Db 657 CTCTGAGTGTGATCTGTTCTGCTCTGCTGATGGCCAGAGGTCCTCAATAAATCTGC 716
 QY 532 CTAGCCTG 539
 Db 717 CTAGCATG 724

RESULT 4
 AAD32057
 ID AAD32057 standard; DNA; 580 BP.
 XX
 AC AAD32057;
 DT 18-JUN-2002 (first entry)
 XX
 DE Pigmy rattlesnake venom gland protein, Zsnk4 gene.
 XX
 XX Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;
 KW platelet aggregation; gene; Zsnk4; ds.
 XX
 OS Sistrurus miliarius.

Key Location/Qualifiers
 CDS 3..437
 FT /*tag= a
 FT /product= "Zsnk4 protein"
 FT /note= "CDS does not include start codon"
 FT /partial
 FT sig_peptide 3..29
 FT /*tag= b
 FT mat_peptide 30..434
 FT /*tag= c
 FT /product= "Mature Zsnk4 protein"

PN WO200214364-A2.

XX 21-FEB-2002.

XX 13-AUG-2001; 2001WO-US25310.

XX 14-AUG-2000; 2000US-225072P.

XX 14-AUG-2000; 2000US-225087P.

XX 15-AUG-2000; 2000US-225489P.

XX 15-AUG-2000; 2000US-225490P.

XX 20-DEC-2000; 2000US-356997P.

XX (ZYMO) ZYMOGENETICS INC.

XX Sheppard PO, Bishop PD;

XX WPI; 2002-269180/31.

XX P-PSDB; AAE20180.

XX New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which

XX affect blood coagulation and platelet aggregation system, useful in

XX therapy and diagnostics, or as tools in the study of genetics or

XX molecular biology

XX Claim 5; Page 75; 79pp; English.

XX The invention relates to new pigmy rattlesnake (Sistrurus miliarius)

XX venom gland proteins, which affect blood coagulation and platelet

XX aggregation system. The polypeptides, which affect blood coagulation and

XX platelet aggregation system are useful in therapy and diagnostics. The

CC practical kits for courses related to genetics and molecular biology,
 CC protein chemistry and antibody production and analysis. The
 CC polynucleotide or polypeptide can be used as standards or as unknowns
 CC for testing purposes. The polypeptides are also useful in identifying
 CC proteins by western blotting, protein purification, determining the
 CC weight of expressed polypeptides as a ratio to total protein expressed,
 CC identifying peptide cleavage sites, coupling amino and carboxyl terminal
 CC tags, mass spectrometry, circular dichroism to determine conformation or
 CC affinity chromatography columns to purify the protein, cloning or
 CC sequencing. The present sequence is Sistrurus miliarius venom gland
 CC protein, Zsnk4 gene.
 XX

SQ Sequence 580 BP; 154 A; 133 C; 163 G; 130 T; 0 other;

Query Match 52.5%; Score 285.8; DB 24; Length 580;
 Best Local Similarity 75.2%; Pred. No. 1.2e-79;
 Matches 413; Conservative 0; Mismatches 97; Indels 39; Gaps 3;

QY 1 GAAGGCAATGCTACAAAGTCTTCAAAACAATTAAGACCTGGACAGATGCAGAGAGCTTC 60
 Db 63 GATCAGTATTGCTACAGGGTCTCAAAACAATCAGGACCTGGACAGATGCAGAGAGCTTC 122
 QY 61 TGCACGAAGCAGTGAACGGGGGCGATCTGCTCTATCGAAGCTCCGAGAGAGCAGAC 120
 Db 123 TGCTGGACAGCGAAGGGCGGCGATCTGCTCTATTGAAGCGACGAGAGAGCAGCC 182
 QY 121 TTTGTGGCCAGTTGATTGCTCAGAAAGATAAAGTACGCCAAATCCATCTCTGGATCGGA 180
 Db 183 TTTGTGGCCAGCTGCTGCTCAGAACATCAAGCAAAACAATATGATCTGGATCGGA 242
 QY 181 CTGAGGGCTCAAAACAAGAAAGCAATGACATAGAGTGGAGGATGGTCCAGCATC 240
 Db 243 CTGAGGATTCAGGGCGAAGAGCAATGACAGCAAGTGGAGGATGGTCCAGCGTC 302
 QY 241 AGTTATGAGAATTGATTGAAGAAGATCCAAAGATGCTTTGGGTGCACATAGAGACA 300
 Db 303 AATTATGAGACCTGATTAAACATCGGACCAAAAGTGTTTGGGCTGAAAAAGAGACA 362
 QY 301 GGGTTTCATAAGTGGGAGAATTTTACTGTGAACAACAAGATCCTTTTGTCTGCGA--- 356
 Db 363 GGGTTTCGACGTGGCGCAATGTTTCACTGTACACAACAATCTTTTCATGTGCAAGTTC 422
 QY 357 --GGCATAGTCTGAGATCCAGCTGATTGAAGTCTGGAGAGCAAGAGAGCCGCCACCC 414
 Db 423 CCGCCAGAGTGTTAAGATCCGGCTGTGTGAAGTCTGGAGAGCAAGAGAGAGCCGCCACCC 482
 QY 415 CATCCCCCAACCTGCCTAGCCACAATCTCTGCTATGCACCTTTGCTCAACGGATGCTC 474
 Db 483 CACCGC-----CACCTTTGCTCAACGGATGCTC 511
 QY 475 TCTGTAGCTGATCTGGTGTGCTGCTCTCTGATGGCGCGAAG--TCAATAAATCTGCGC 532
 Db 512 TCTGTAGCTGATCTGGTGTGCTGCTCTCTGATGGCGCGAAGTCCCAATAAATCTTCC 571
 QY 533 TAGCCTGAA 541
 Db 572 TAGCCTGGA 580

RESULT 5

AAL42015

ID AAL42015 standard; cDNA; 632 BP.

XX AC AAL42015;

XX 16-MAY-2002 (first entry)

XX Korean adder snake venom salmorin A chain protein cDNA sequence.

XX Korean adder; ss; salmorin protein A chain; snake venom;

XX fibrinogen clotting inhibition; thrombosis; prothrombin binding;

XX thrombin binding; blood coagulation.

XX

```

OS Agkistrodon halyx brevicaudus.
XX
FH Key Location/Qualifiers
FT CDS 34..498
FT /*tag= a
FT /product= "Salmonin A chain protein"
FT sig_peptide 34..102
FT /*tag= b
FT /note= "Signal peptide"
FT mat_peptide 103..495
FT /*tag= c
FT /note= "Mature Salmonin A chain protein"
FT 3'UTR 499..622
FT /*tag= d
FT polyA_signal 623..632
FT /*tag= e
XX
PN WO200214514-A1.
XX
XX 21-FEB-2002.
XX
XX 26-JUL-2001; 2001WO-KR01277.
XX
XX 27-JUL-2000; 2000KR-0043470.
XX
XX (BIOB-) BIOUD CO LTD.
XX
XX Chung K, Kim D, Koh Y;
XX
XX WPI; 2002-241907/29.
XX P-PSDB; AAO14520.
XX
XX New salmonin protein derived from venom of Korean adder Agkistrodon
XX halyx brevicaudus, useful for treating thrombosis by repressing
XX fibrinogen clotting through repression of activation of prothrombin
XX into thrombin -
XX
XX Claim 1; Fig 1A; 30pp; English.
XX
XX The invention comprises the nucleotide and protein sequences of a
XX salmonin protein derived from the venom of Korean adder. Salmonin protein
XX is composed of an A chain and a B chain, and has inhibitory activity
XX against fibrinogen clotting. The salmonin protein of the invention is
XX useful for treating thrombosis, as it represses fibrinogen clotting
XX potentially by binding to prothrombin and thrombin so as to delay blood
XX coagulation. The present cDNA sequence encodes the Korean adder salmonin
XX A chain protein.
XX
XX Sequence 632 BP; 162 A; 140 C; 182 G; 148 T; 0 other;
XX
XX Query Match 50.8%; Score 276.2; DB 24; Length 632;
XX Best Local Similarity 80.0%; Pred. No. 1.4e-76;
XX Matches 337; Conservative 0; Mismatches 83; Indels 1; Gaps 1;
XX
QY 2 AAGGCATCTACAGGCTTCAACATCTAAGACCTGACAGATGACAGAGCTTCT 61
Db 137 ATGGTCAATGCTACAGGCTTCAACATCTAAGACCTGACAGATGACAGAGCTTCT 196
QY 62 GCACCAAGCAGTGAACGGGGGCGATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGACT 121
Db 197 GTCAGCAGCGGAGGCGGGGCGATCTGGTCTCTATCGAAACACGGGAGAGCAGACT 256
QY 122 TTGTGGCCAGTTGATGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCGATCGGAC 181
Db 257 TTGTGGCCACGCTGGTCTGAGAGGATAGAGACATCTTCCCATGCTGATCGGAC 316
QY 182 TGAGGCTCAAAACAAAGAAAGCAATGACATAGAGTGGAGCGATGCTCAGATCA 241
Db 317 TGAGGATCAAGGCAAGCAAGCAATGACATGAGTCTGAGTGGAGCGATGATCCAGGTCA 376
QY 242 GTTATCAGATTCGATTGAAGAGATCCAAAAGTCTCTGGGGTGACATAGACAG 301
Db 377 GTTATGAGAACTGGATTGAAGAGATCCAAAAGTCTCTGGGGTGAGACTCA 436

```

```

QY 302 GGTTCATAAGTGGAGAAATTTTACTGTGAAACAAGAATCCTTTTGTCTGCGAGCAT 361
Db 437 ATTATCATAAGTGGTCAATGTTTACTGTGCAAAAGAAATCCTTTGCTGCGAGGCTT 496
QY 362 AGT-CTGAAGATCCAGCTGATTGAAGTCTGAGAAGCAAGAGCCGCCACCCATCCC 420
Db 497 AGTCTGAAGATGAGCTGTGTGAAGTCTGAGAAGCAAGAGCCGCCACCCATCCG 556
QY 421 C 421
Db 557 C 557

RESULT 6
AAI71876
ID AAI71876 standard; DNA; 601 BP.
XX
XX AAI71876;
XX
XX 10-JAN-2002 (first entry)
XX
XX Snake venom blood anticoagulant halyxin A chain coding sequence.
XX
XX Snake; pit viper; venom; halyxin; anticoagulant; thrombogenesis;
XX thrombosis; ds.
XX
XX Agkistrodon halyx.
XX
XX KR2001049671-A.
XX
XX 15-JUN-2001.
XX
XX 29-JUN-2000; 2000KR-0036591.
XX
XX 29-JUN-1999; 99KR-0025105.
XX
XX (BIOB-) BIOUD CO LTD.
XX
XX Jang YS, Jung GH, Kim DS, Koo BH, Son YD;
XX
XX WPI; 2001-637330/73.
XX P-PSDB; AAM51543.
XX
XX Halyxin as blood anticoagulation protein separated from snake venom -
XX Claim 1; Page 9; 21pp; Korean.
XX
XX The invention relates to halyxin, a novel protein with very strong
XX blood anticoagulation activity. The protein was separated from snake
XX venom of Agkistrodon halyx brevicaudus (a Korean pit viper) and can
XX be used in the treatment of thrombogenesis. The present sequence
XX encodes the A chain of halyxin.
XX
XX Sequence 601 BP; 166 A; 133 C; 163 G; 139 T; 0 other;
XX
XX Query Match 48.9%; Score 266; DB 22; Length 601;
XX Best Local Similarity 82.4%; Pred. No. 2.2e-73;
XX Matches 305; Conservative 0; Mismatches 65; Indels 0; Gaps 0;
XX
QY 1 GNAGGGCATTCGTAAGGCTTCAACATCTAAGACCTGACAGATGACAGAGCTTC 60
Db 120 GAAGGGCATTCGTAACATCTTCCATCTATTCAAGACCTGGCAGAGCAGAGAGTTC 179
QY 61 TCACAGAGCAGGTGAACGGGGGCGATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
Db 180 TCACAGAGCAGGTGAAGGGCGCGATCTGGTCTCTATCGAAAGCTCCGAGAAGCAGAC 239
QY 121 TTTGTGGGCGAGTTGATGCTCAGAAGATAAAGTCAAGCAAAATCCATGCTCGGATCGGA 180
Db 240 TTTGTGGGCGAGTGGTCTCTGAGAACAATGAAGACATCGGATCTATATCTGGATCGGA 299
QY 181 CTGAGGGCTCAAAACAAGAAAGCAATCCACATAGATGGAGCGATGGCTCCAGCATC 240

```

Db 300 CTGAGGGTTCGAGCAAAAGAGCAATGCAGCTCCAGTGGAGCGATGCTCCAGGTC 359
 QY 241 AGTTATCAGAAATTCGAGTGAAGAGCAATCAAAAAGTGTCTGGGTGCAATAGAGACA 300
 Db 360 AGTTATCAGAAATTCGAGTGAAGAGCAATCAAAAAGTGTCTGGGTGCAAAAAGAGACA 419
 QY 301 GGGTTTCATAAGTGGGAGAAATTTTACTGTGAACAACAAGATCCTTTGTCTCGGAGGCA 360
 Db 420 GAGTTTCGTAAGTGGTTCATATTTACTGTGGAGAAAGCAATCCTTTCGCTCGGAGGCA 479
 QY 361 TAGTCTGAAG 370
 Db 480 TAGAGCAAG 489

RESULT 7

AAQ89309
 ID AAQ89309 standard; cDNA; 690 BP.

AC AAQ89309;

DT 25-MAR-2003 (updated)

DT 28-NOV-1995 (first entry)

XX Snake venom antithrombotic oligopeptide cDNA.

KW Antithrombotic peptide; snake venom; platelet binding inhibition;
 von Willebrand factors; Crotalus horridus horridus; ds.

OS Crotalus horridus horridus.

FH Key Location/Qualifiers
 FT CDS 66..515
 FT /*tag= a

PN WO9508573-Al.

XX 30-MAR-1995.

PF 21-SEP-1994; 94WO-JP01555.

PR 22-SEP-1993; 93JP-0236975.

XX (AJIN) AJINOMOTO KK.

PI Fukuchi N, Ishii K, Kito M, Kobayashi T, Nagano M;
 Tanaka A, Yamamoto H, Yoshimoto R;

DR WPI; 1995-139559/18.

DR P-PSDB; AAR71981.

XX Single-chain antithrombotic peptide - obtained by cleaving an
 oligopeptide from snake venom to break inter-chain di: sulphide
 bonds but preserve intra-chain di: sulphide bonds

PS Example 2; Pages 47-48; 84pp; Japanese.

XX AAR71978 and AAR71979 are snake venom derived antithrombotic peptides,
 specifically from the snake venom oligopeptide AAR71981, encoded by
 AAQ89309. These peptides have the advantage of avoiding significant
 thrombocytopenia when administered at the minimum dose, for in vivo
 inhibition of platelet von Willebrand factor binding.
 (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 690 BP; 167 A; 180 C; 170 G; 173 T; 0 other;

Query Match 41.6%; Score 226.4; DB 16; Length 690;
 Best Local Similarity 70.5%; Pred. No. 7.2e-61;
 Matches 389; Conservative 0; Mismatches 126; Indels 37; Gaps 5;

QY 1 GAAGGGCATTCCTCAAGGTCTTCAACAATCTAAGACCTGACAGATGCGAGAGCTTC 60

Db 168 GATCGGTATTGCTACAGCCCTTCAAAACAAGAGATGACCTGGGCCGATGAGAGAGGTTTC 227
 QY 61 TGCACGAAGCAGGTGAACCGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAGACAGAC 120
 Db 228 TGCTCGGAGCAGCGAAGCGGGCATCTCTCTGTGCGAAACCGCCCTAGAGCATCC 287
 QY 121 TTTGTGGGCAGTGTGATTGCTCAGAAGATAAAGTCAAGCAAAATCCATGCTCGATCGGA 180
 Db 288 TTTGTGGCAATGTGCTCTATGCGAAACAAGAGTACCTCACAGTTATATCTGGAATTGGA 347
 QY 181 CTGAGGGCTCAAAAACAAGAAAGCAATGCAGCATAGATGGAGCGATGCTCCAGCATC 240
 Db 348 CTGAGGGTCAAAAACAAGAGCAGCCAT-----GCTCCAGCATC 386
 QY 241 AGTTATGAGAAATTCGAAAGAAGAAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACA 300
 Db 387 AGTTATGAGAACTCTGGTT-----GACCCATTTGAATGTTTATGTGAGCAGAGACACA 440
 QY 301 GGGTTTCATAAGTGGGAGAAATTTTACTGTGAACAACAAGATCCTTTGTCTGCGA---- 356
 Db 441 AGCTTCGTGAGTGGTTTAAAGTGTGACTGTGAACAACAACATCTTTTATATGCAAGTTTC 500
 QY 357 --GGCATAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGAGCCGCCACCC 414
 Db 501 ACGCGACCACGTTAGATCCGCTGTGTGAAGTCTGGAGAAGCAAGAGCCGCCACCT 560
 QY 415 CATCCCCCAACCTGCTAGCCACCAATCTCTGCTATGCAACCTTTGCTCAAGGATGCTC 474
 Db 561 C--TCCCCACCCGCCACCTTCGCAATCTCTGCTCTCTCCCTTTGCTCAGTGGATGCTC 618
 QY 475 TCTGTAGTGGATCTGGTGTGCTGCTCTCTGTGAGTGGCCGGAAG--TCAATAAATTCGTC 532
 Db 619 TCTGTAGCCGATCTGGGTTTCTGCTCAGATGGGTGAGAGATCCATAAATTCGTC 678
 QY 533 TAGCCTGAAAA 544
 Db 679 TACCCAAAAA 690

RESULT 8

AAC61144

ID AAC61144 standard; DNA; 690 BP.

AC AAC61144;

DT 07-FEB-2001 (first entry)

DE DNA encoding a snake venom derived protein.

KW Subunit peptide production; snake venom; rattlesnake; thrombolytic;
 von Willebrand's factor; blood platelet-inhibitory activity; ds.

OS Crotalus horridus horridus.

XX WO200059926-Al.

XX 12-OCT-2000.

PF 31-MAR-2000; 2000WO-JP02127.

PR 02-APR-1999; 99JP-0096073.

PA (AJIN) AJINOMOTO CO INC.

XX Fukuchi N, Kageyama S, Kito M, Kayahara T, Yamamoto H;

XX WPI; 2000-664985/64.

DR P-PSDB; AAY85628.

PT Producing physiologically-active subunit peptides originating in
 polymer proteins by denaturation and specific separation, with lower
 antigenicity but improved solubility and stability, e.g. blood
 platelet-binding inhibitors

XX Disclosure; Page 46; 51pp; Japanese.

XX This invention relates to a method for the production of a subunit

XX peptide originating from a polymer protein with disulphide bonds within

CC and between subunits. The method comprises denaturing the protein or its

CC subunit using a protein denaturing agent in a solution, removing the

CC agent in the presence of a polyoxyalkyl polyether which reacts with a

CC thiol group and unwinds the subunit, and separating the polyoxyalkyl

CC polyether-bound subunit peptide. The method can be used for producing

CC physiologically-active subunit peptides for polymer proteins e.g. snake

CC venom-originated dimer peptide with blood platelet-inhibitory activity

CC on von Willebrand's factor. The peptides produced have platelet-binding

CC inhibitory, and thrombolytic activity. The present sequence represents

CC DNA encoding rattlesnake protein used in an example illustrating the

CC method of the invention.

XX Sequence 690 BP; 167 A; 180 C; 170 G; 173 T; 0 other;

Query Match 41.6%; Score 226.4; DB 21; Length 690;
Best Local Similarity 70.5%; Pred. No. 7.2e-61;
Matches 389; Conservative 0; Mismatches 126; Indels 37; Gaps 5;

QY 1 GAAGGGCATTTGCTACAGGTCTTCAACAATCTAAGACCTCGGCAGATCGAGAGCTTC 60
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 168 GATCGGTATTGCTACAGCCCTTCAAACAGAGATGACCTGGGCCGATCGCAGAGGTT 227
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 61 TGCCAGAAGCAGGTGAACGGGGGGCATCTGGTGCTATPCGAAAGTCCCGAGAAGCAGAC 120
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 228 TGCTCGNGCAGCGGAGGGCGGCNCTCTCTCTGTCGAAACGGCCCTAGAGCATCC 287
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 121 TTTGTGGGCCAGTTGATGTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 288 TTTGTGACAAATGTGCTCTATGCGAAACAAAGAGTACCTCACACGTTATATCTGGATTGGA 347
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 181 CTGAGGGCTCAAAACAAAGAAAACCAATGCAGCATAGAGTGGAGCGGATGGCTCCAGCATC 240
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 348 CTGAGGGTTCAAAAACAAAGGACGCCAT-----GCTCCAGCATC 386
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 241 AGTTATGAGAAATTCGATTGAAGAAGAAATCCAAAAAGTGTCTGGGTGCACATAGAGACA 300
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 387 AGTTATGAGAACCTGGTT-----GACCATTGAAUGTTTTATGTGAGCAGAGACACA 440
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 301 GGGTTTTCATAAGTGGGAGAAATTTTACTGTGAACAAACAAGATCCCTTTTGTCTGCGA --- 356
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 441 AGGCTTCGTGAGTGGTTTAAAGTTGACTGTGAACAACAACATCTTTCATATGCAAGTTC 500
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 357 --GGCATAGTCTGAAGATCCAGCTGATTTGAAGTCTGGAGAAGCAAGAGCCCCCCCACCC 414
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 501 ACGGGACCCAGCTTAAGATCCGGCTGTGTGAAGTCTGGAGAAGCAGGAAGAGCCCCCACCT 560
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 415 CATCCCCAACCTGCCTPAGCCAAATCTCTGTGATGACCCCTTTGTCTCAACGGATGCTC 474
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 561 C--TCCCCACCCCCACCTTCGGAAATCTGTCTTCCCCCTTTGCTAGTGGATGCTC 618
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 475 TCTGTAGCTGGATCTGGTGTGTGCTGCTCCTGATGGCCGGGAAG--TCAATAAATCTGCC 532
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 619 TCTGTAGCCGATCTGGGTTTTCTGCTCCAGATGGTCTCAGAAGATCCAATAAATCTGCC 678
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 533 TAGCCTGAAAAA 544
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 679 TACCCAAAAAAA 690
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 9
AAK99834
ID AAK99834 standard; DNA; 690 BP.
XX AC AAK99834;
XX XX
DT 19-JUL-2002 (first entry)
XX DE
DNA encoding the antithrombotic wild-type rattlesnake protein.

Db 501 ACGGACACCGATTAGATCCGGCTGTGTGAAGTCTGGAGAAGCAAGGAAGCCCCACCT 560
 QY 415 CATCCCCACACCTGCTAGCCACAACTCTGTATGACACCTTTGCTCAACGATGCTC 474
 Db 561 C--TCCCCACCCCCACCTTCCGCAATCTGTCTCTTCCCCCTTTGCTCAGTGGATGCTC 618
 QY 475 TCTGTAGCTGGATCTGGTGTGTGCTCTCTGTATGAGGCGCGGAAG--TCAATAAATTTCTGCC 532
 Db 619 TCTGTAGCGGATCTGGTGTGTGCTCTCAGATGGGTGAGAGATCAATAAATTTCTGCC 678
 QY 533 TAGCCTGAATAA 544
 Db 679 TACCCAAAAA 690

RESULT 10

AAD32056
 ID AAD32056 standard; DNA; 456 BP.

AC AAD32056;

XX 18-JUN-2002 (first entry)

DE Pigmy rattlesnake venom gland protein Zsnk3, degenerate nucleic acid.

KW Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;

KW platelet aggregation; ds.

XX Sistrurus miliarius.

XX WO200214364-A2.

XX 21-FEB-2002.

XX 13-AUG-2001; 2001WO-US25310.

XX 14-AUG-2000; 2000US-225072P.

XX 14-AUG-2000; 2000US-225087P.

XX 15-AUG-2000; 2000US-225489P.

XX 15-AUG-2000; 2000US-225490P.

XX 20-DEC-2000; 2000US-356997P.

XX (ZYMO) ZYMOGENETICS INC.

XX Sheppard PO, Bishop PD;

XX WPI; 2002-269180/31.

XX P-PSDB; AAE20179.

XX New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which affect blood coagulation and platelet aggregation system, useful in therapy and diagnostics, or as tools in the study of genetics or molecular biology -

XX Disclosure; Page 74; 79pp; English.

XX The invention relates to new pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which affect blood coagulation and platelet aggregation system. The polypeptides, which affect blood coagulation and platelet aggregation system, are useful in therapy and diagnostics. The polypeptides are also useful as an educational tool in laboratory practical kits for courses related to genetics and molecular biology, protein chemistry and antibody production and analysis. The polynucleotide or polypeptide can be used as standards or as unknowns for testing purposes. The polypeptides are also useful in identifying proteins by western blotting, protein purification, determining the weight of expressed polypeptides as a ratio to total protein expressed, identifying peptide cleavage sites, coupling amino and carboxyl terminal tags, mass spectrometry, circular dichroism to determine conformation or affinity chromatography columns to purify the protein, cloning or sequencing. The present sequence is Sistrurus miliarius venom gland protein Zsnk3, degenerate nucleic acid.

SQ Sequence 456 BP; 72 A; 31 C; 92 G; 66 T; 195 other;

Query Match 34.3%; Score 186.6; DB 24; Length 456;

Best Local Similarity 47.1%; Pred. No. 2.1e-48;

Matches 169; Conservative 86; Mismatches 104; Indels 0; Gaps 0;

QY 1 GAAGGGCATTGCTACAGGCTTTCAACAATCTTAAGACCTGGACAGATGAGAGCTTC 60

Db 97 GAYCARCAYTGTAYMGNGTNTTYAARCARCTYNAARACNTGGGAYGAYGNGARMGNTTY 156

QY 61 TGCACGAAGCAGTGAACGGGGGCATCTGCTCTATCGAAAGCTCCGAGAGCAGAC 120

Db 157 TGWNSGARCARGCARGGNGNCAYYNGTNWNSNATHGARWNSNGARGARGNGCN 216

QY 121 TTTGTGGCCAGTTGATCTCTCAGAGATAAAGTACGCCAAATCCATCTCTGGATCGGA 180

Db 217 TTYGTNGCNCARYTNGTNCNGARAAYMGNGMNGCNATHYTNATYATHTGATHGNN 276

QY 181 CTGAGGGCTCAAAACAAAGAAAGCAATGCAGATAGATGGAGCGATGGCTCCAGCATC 240

Db 277 YTMGNGTNCARGGNAARGAARCAARTGYNSGNCNAARTGOWNSGAYGNGNSWNGTN 336

QY 241 AGTTATGAGATTTGGATTCAAGAAGAAATCCAAAAGTGTCTTGGGGTGACATAGAGACA 300

Db 337 WNTAYGAAAYTGGATHGARGCNGARWNSAARACNTGYTNGNTNCARCGGNGCN 396

QY 301 GGGTTTCATAAGTGGGAGAAATTTTACTGTGAAACAAAGATCCTTTTCTGCGAGGC 359

Db 397 AAYTAYCAVAARTGGGTAAAYATHATYTGCGNGARATHAAYCCNTTYGNTGYGARGC 455

RESULT 11

AAD32058

ID AAD32058 standard; DNA; 432 BP.

XX AAD32058;

DT 18-JUN-2002 (first entry)

XX Pigmy rattlesnake venom gland protein Zsnk4, degenerate nucleic acid.

XX Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;

XX platelet aggregation; ds.

XX Sistrurus miliarius.

XX WO200214364-A2.

XX 21-FEB-2002.

XX 13-AUG-2001; 2001WO-US25310.

XX 14-AUG-2000; 2000US-225072P.

XX 14-AUG-2000; 2000US-225087P.

XX 15-AUG-2000; 2000US-225489P.

XX 15-AUG-2000; 2000US-225490P.

XX 20-DEC-2000; 2000US-356997P.

XX (ZYMO) ZYMOGENETICS INC.

XX Sheppard PO, Bishop PD;

XX WPI; 2002-269180/31.

XX New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which affect blood coagulation and platelet aggregation system, useful in therapy and diagnostics, or as tools in the study of genetics or molecular biology -

XX Disclosure; Page 76; 79pp; English.

XX The invention relates to new pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which affect blood coagulation and platelet


```

XX OS Agkistrodon halys.
XX PN KR2001049671-A.
XX PD 15-JUN-2001.
XX PF 29-JUN-2000; 2000KR-0036591.
XX PR 29-JUN-1999; 99KR-0025105.
XX PA (BIOB-) BIOBUD CO LTD.
XX PI Jang YS, Jung GH, Kim DS, Koo BH, Son YD;
XX DR WPI: 2001-637330/73.
XX DR P-PSDB; AAM51544.
XX PT Halyxin as blood anticoagulation protein separated from snake venom -
XX PS Claim 1; Page 11; 21pp; Korean.
XX CC The invention relates to halyxin, a novel protein with very strong
XX CC blood anticoagulation activity. The protein was separated from snake
XX CC venom of Agkistrodon halys brevicaudus (a Korean pit viper) and can
XX CC be used in the treatment of thrombogenesis. The present sequence
XX CC encodes the B chain of halyxin.
XX SQ Sequence 633 BP; 179 A; 142 C; 157 G; 155 T; 0 other;

Query Match 26.5%; Score 144; DB 22; Length 633;
Best Local Similarity 63.4%; Pred. No. 6.7e-35;
Matches 350; Conservative 0; Mismatches 140; Indels 62; Gaps 6;

QY 1 GAAGGGCATGCTACAAAGTCTTCAACAACTTAAGACCTGACAGATGACAGAGCTTC 60
DB 120 GAAGGGCATGCTACAAAGTCTTCAACAACTTAAGACCTGACAGATGACAGAGCTTC 179
QY 61 TGCACGAAGAGGTGAACGGGGGCTCTGCTCTATCGAAGCTCCGGAGAAGCAGAC 120
DB 180 TGCACACACAGCAGCAGAGGGGATCTGCTCTCTTCCACGACCTGAAGAGCAGAT 239
QY 121 TTTGTGGCCAGTTGATGCTCAGAAAGATAAGTCAGCCAAATCCATGCTGTGATCGGA 180
DB 240 TTTGTGTGCAAGCTGGCCCTTCCA-----AAATTTGGCCAGGTATTTCTGTGATGGA 293
QY 181 CTGAGGGCTCAAAACAAAGAAAGCAATGCAGCATAGATGGAGCGATGGCTCCAGATC 240
DB 294 CTGAG-----CAATGCTGGAATCAATGCAGCTGGCAATGGAGCAGTGTGCCAAGCTC 347
QY 241 AGTTATGAGAATTGGATTGAAGAAGAAATCCAAAGATGCTTTGGGGTGCACATAGAGACA 300
DB 348 AAATACGAAGCTGGCTGAAGAA-----TCTTATGTGCTATTTCAAGTCA 395
QY 301 GGGTTTCAATAGTGGAGAAATTTTATCTGTGAACAAACAAAGATCCCTTTTGTCTGG- 355
DB 396 ACTAATACAAATAGGAGAGTAGAGCTCTGCAGATGGAGGCATATTTCTGTGGAGTTC 455
QY 356 -AGCATAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAGCAGAGAGCCGCCACCC 414
DB 456 CAGCATAGTTTGAAGATCCAGCTGAGTGAAGTCTGGAGAGCAAGGAAGCCGCCACAT 515
QY 415 CATCCCCCAACCCCTGCCCTAGCCACAACTCTGTGCTATGCACCTTTTGTCTCAACGGATGCTC 474
DB 516 CACCCCC-----CACCTTCGCTCAATGATGCTC 545
QY 475 TCTGTAGCTGATCTGGTGTGCTGCTCTCTGATGGCCGGAAG--TCAATAAATTTCTGCC 532
DB 546 TCTGTAGCTGATCTGGTGTGCTGCTCTCTGATGGCCGGAAGTCCCAATAAATTTCTGCC 605
QY 533 TAGCCTGAAAAA 544
DB 606 TAGCATGAAAAA 617

```

```

RESULT 14
AAQ9308
ID AAQ9308 standard; cDNA; 272 BP.
XX AC AAQ9308;
XX DT 25-MAR-2003 (updated)
XX DT 28-NOV-1995 (first entry)
XX DE Snake venom antithrombotic oligopeptide cDNA.
XX KW Antithrombotic peptide; snake venom; platelet binding inhibition;
XX KW von Willebrand factors; Crotalus horridus horridus; ds.
XX OS Crotalus horridus horridus.
XX PN WO9508573-A1.
XX PD 30-MAR-1995.
XX PF 21-SEP-1994; 94WO-JP01555.
XX PR 22-SEP-1993; 93JP-0236975.
XX PA (AJIN) AJINOMOTO KK.
XX PI Fukuchi N, Ishii K, Kito M, Kobayashi T, Nagano M;
XX PI Tanaka A, Yamamoto H, Yoshimoto R;
XX DR WPI: 1995-139559/18.
XX PT Single-chain antithrombotic peptide - obtained by cleaving an
XX PT oligopeptide from snake venom to break inter-chain di: sulphide
XX PT bonds but preserve intra-chain di: sulphide bonds
XX PS Example 2; Page 46; 84pp; Japanese.
XX CC AAQ9308 is a snake venom cDNA used in the prepn. of AAQ9309, which
XX CC encodes AAR71981 a snake venom oligopeptide. From this oligopeptide
XX CC claimed antithrombotic peptides were derived, the peptides had the
XX CC advantage of avoiding significant thrombocytopenia when administered
XX CC at the minimum dose, for in vivo inhibition of platelet von Willebrand
XX CC factor binding.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 272 BP; 72 A; 60 C; 77 G; 63 T; 0 other;

```

```

Query Match 18.9%; Score 102.6; DB 16; Length 272;
Best Local Similarity 72.9%; Pred. No. 5e-22;
Matches 132; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 34 AAGACCTGGACAGATGACAGAGCTTCTGCACGAAGCAGGTGAACGGGGGCTCTGGTC 93
DB 7 ATGACTTGGCCGATGACAGAGGTTCCTCGGAGCAGCGGCGGCGATCTCTCTC 66
QY 94 TCTATCGAACTCCGGAGAAGCAGACTTTTGTGGCCAGTTGCTTCAGAAGATAAG 153
DB 67 TCTGTGAAACCGCCCTAGAGCATCTCTTTGTGGACAATGTGCTCTATGCGAACAAAGAG 126
QY 154 TCAGCCAAATCCATGCTGTGATCGAGTGGCTCAAAACAAAGAAAGCAATGCGAGC 213
DB 127 TACTTCACAGTTATATCTGGATTGGAGTGGGTTTCAAAACAAAGGACAGCCATGCTCC 186
QY 214 A 214
DB 187 A 187

```

```

RESULT 15
AAC61143
ID AAC61143 standard; DNA; 272 BP.

```


This Page Blank (uspto)

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2003, 10:25:43 ; Search time 79 Seconds
(without alignments)
3039.398 Million cell updates/sec

Title: US-09-938-114-1

Perfect score: 544
Sequence: 1 GAGGGCAATGCTACAGGT.....ATTCTGCCTAGCCTGAAAAA 544

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/5A COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B COMB.seq: *
3: /cgn2_6/ptodata/1/ina/5A COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PTCUS COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	544	100.0	544	4	US-09-058-740-1
2	226.4	41.6	690	2	US-08-612-840A-7
3	102.6	18.9	272	2	US-08-612-840A-6
4	58.4	10.7	7218	1	US-08-232-463-14
5	58	10.7	454	4	US-09-058-740-12
6	34.6	6.4	2949	4	US-09-412-554A-3
7	34.2	6.3	590	2	US-08-454-557C-32
8	34.2	6.3	590	2	US-08-340-426D-32
9	34.2	6.3	590	2	US-08-450-673C-32
10	34.2	6.3	590	5	PCT-US95-17111A-32
11	32.6	6.0	777	3	US-09-146-969-1
12	32.4	6.0	614	2	US-08-729-103-2
13	32.4	6.0	1114	2	US-08-468-413-1
14	32.4	6.0	1114	3	US-09-162-508-1
15	32.4	6.0	1114	5	PCT-US95-07169-1
16	32.2	5.9	562	4	US-09-669-751-15
17	32	5.9	1201	4	US-09-461-325-42
18	31.6	5.8	531	1	US-08-340-539A-5
19	31.6	5.8	531	2	US-08-461-592B-5
20	31.6	5.8	1829	6	5514582-1
21	31.6	5.8	2359	2	US-08-513-278-1
22	31.6	5.8	2330	1	US-08-481-803-1
23	31.6	5.8	2330	1	US-08-215-366A-1
24	31.6	5.8	2330	1	US-08-340-539A-1
25	31.6	5.8	2330	2	US-08-461-592B-1
26	31.6	5.8	2385	4	US-09-495-050A-292
27	31.6	5.8	5169	4	US-09-194-612A-2

28 31.4 5.8 8906 2 US-08-826-267-1 Sequence 1, Appli
29 31.2 5.7 46819 4 US-09-453-702B-72 Sequence 72, Appli
c 30 30.8 5.7 5261 1 US-08-045-806-3 Sequence 3, Appli
31 30.8 5.7 5261 1 US-08-366-051B-3 Sequence 3, Appli
c 32 30.6 5.6 280 6 5198345-11 Patent No. 5198345
33 30.6 5.6 289 3 US-09-007-005-17 Sequence 17, Appli
34 30.6 5.6 289 3 US-09-244-796-17 Sequence 17, Appli
c 35 30.6 5.6 289 6 5198345-18 Patent No. 5198345
36 30.4 5.6 747 2 US-08-401-530A-1 Sequence 1, Appli
37 30.4 5.6 2325 2 US-08-714-677-3 Sequence 3, Appli
38 30.4 5.6 2325 2 US-08-393-540-3 Sequence 3, Appli
39 30.4 5.6 2325 2 US-08-714-537-3 Sequence 3, Appli
40 30.4 5.6 3073 2 US-08-714-677-11 Sequence 11, Appli
41 30.4 5.6 3073 2 US-08-393-540-11 Sequence 11, Appli
42 30.4 5.6 3073 2 US-08-714-537-11 Sequence 11, Appli
43 30.2 5.6 762 2 US-08-822-261-2 Sequence 2, Appli
44 30.2 5.6 762 4 US-09-226-852-2 Sequence 2, Appli
c 45 30 5.5 406 4 US-09-556-877-50 Sequence 50, Appli

ALIGNMENTS

RESULT 1
US-09-058-740-1
; Sequence 1, Application US/09058740
; Patent No. 6489451
; GENERAL INFORMATION:
; APPLICANT: Chun Wang, Benjamin XY Li, Xin Cheng,
; Jing Liu, Li-Wen Niu, Wan-Zhi Huang,
; Zhen-Yu Xu, Dan Luo, Lian-Di Kang,
; Jin-Guo Ding, Fang Rong, Yan Liu and
; Hui-Ran Chen
; TITLE OF INVENTION: AN ANTITHROMBOSIS ENZYME FROM THE SNAKE
; VENOM OF AGKISTRODON ACUTUS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; FILING DATE: 10-Apr-1998
; APPLICATION NUMBER: US/09/058,740
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Chen, Anthony C.
; REGISTRATION NUMBER: 38,673
; REFERENCE/DOCKET NUMBER: 233/298
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 544 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-058-740-1

```
Query Match      100.0%; Score 544; DB 4; Length 544;
Best Local Similarity 100.0%; Pred. No. 1.3e-164;
Matches 544; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAGGGCATTGTCTACAAAGTCTTCAAAACAATTAAGACCTTGACAGATCGACAGAGCTTC 60
DB 1 GAAGGGCATTGTCTACAAAGTCTTCAAAACAATTAAGACCTTGACAGATCGACAGAGCTTC 60

QY 61 TGCACGAAGCAGGTGAACGGGGGCGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
DB 61 TGCACGAAGCAGGTGAACGGGGGCGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120

QY 121 TTTGTGGGCGAGTGTGATCTCAGAGATAAAGTCAGCCAAATCCATCTCTGGATCGGA 180
DB 121 TTTGTGGGCGAGTGTGATCTCAGAGATAAAGTCAGCCAAATCCATCTCTGGATCGGA 180

QY 181 CTGAGGGCTCAAAACAAAGAAAGCAATGCAGATAGATGGAGCGATGGCTCCAGCATC 240
DB 181 CTGAGGGCTCAAAACAAAGAAAGCAATGCAGATAGATGGAGCGATGGCTCCAGCATC 240

QY 241 AGTTATGAGAAATGATTAAGAGAAATCCAAAAGTGTCTTGGGGTGACATAGAGACA 300
DB 241 AGTTATGAGAAATGATTAAGAGAAATCCAAAAGTGTCTTGGGGTGACATAGAGACA 300

QY 301 GGGTTTCATAAGTGGGAGAAATTTTACTGTGAACAACAGATCCTTTTGTCTGGAGGCA 360
DB 301 GGGTTTCATAAGTGGGAGAAATTTTACTGTGAACAACAGATCCTTTTGTCTGGAGGCA 360

QY 361 TAGTCTGAAGATCCAGTCTGATTAAGTCTGGAAGCAAGGAGCCGCCATCC 420
DB 361 TAGTCTGAAGATCCAGTCTGATTAAGTCTGGAAGCAAGGAGCCGCCATCC 420

QY 421 CCAACCTCCCTAGACCAATCTCTGATGACACCTTTGCTCAACGAGATGCTCTCTGTA 480
DB 421 CCAACCTCCCTAGACCAATCTCTGATGACACCTTTGCTCAACGAGATGCTCTCTGTA 480

QY 481 GCTGATCTGGTGTGCTCTCTGATGGCGGAGTCAATAAATCTGCTAGCCTCA 540
DB 481 GCTGATCTGGTGTGCTCTCTGATGGCGGAGTCAATAAATCTGCTAGCCTCA 540

QY 541 AAAA 544
DB 541 AAAA 544
```

RESULT 2

US-08-612-840A-7
Sequence 7, Application US/08612840A
Patent No. 5856126

GENERAL INFORMATION:

APPLICANT: FUKUCHI, Naoyuki
APPLICANT: YAMAMOTO, Hiroshi
APPLICANT: NAGANO, Mitsuho
APPLICANT: KITO, Morikazu
APPLICANT: TANAKA, Akiko
APPLICANT: ISHII, Koichi
APPLICANT: KOBAYASHI, Tsuyoshi
APPLICANT: YOSHIMOTO, Ryota
TITLE OF INVENTION: PEPTIDE HAVING ANTI-THROMBUS ACTIVITY AND
METHOD OF PRODUCING THE SAME
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:

ADDRESSEE: Ohlon, Spivak, McClelland, Maier & Neustadt, P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: VA

COUNTRY: USA

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,840A
FILING DATE: 20-MAR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-236975
FILING DATE: 22-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: NO. 5856126man F. Oblon
REGISTRATION NUMBER: 24,618
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 690 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Crotales horridus
STRAIN:
FEATURE:
NAME/KEY: CDS
LOCATION: 66..512
US-08-612-840A-7

Query Match 41.6%; Score 226.4; DB 2; Length 690;

Best Local Similarity 70.5%; Pred. No. 8.8e-63;

Matches 389; Conservative 0; Mismatches 126; Indels 37; Gaps 5;

```
QY 1 GAAGGGCATTGTCTACAAAGTCTTCAAAACAATTAAGACCTTGACAGATCGACAGAGCTTC 60
DB 168 GATCGGTTATGCTACAGCCCTTCAACAGAGATGACCTGGGCCGATCGACAGAGCTTC 227

QY 61 TGCACGAAGCAGGTGAACGGGGGCGCATCTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
DB 228 TGCTCGAGCAGCGAAGGGGGCGCATCTCTCTCTGTCGAAACCGCCCTAGAGCATCC 287

QY 121 TTTGTGGGCGAGTGTGATCTCAGAGATAAAGTCAGCCAAATCCATCTCTGGATCGGA 180
DB 288 TTTGTGGCAATGTCTCTATGCGAACAAAGAGTACCTCACACGTTATATCTGGATTGA 347

QY 181 CTGAGGGCTCAAAACAAAGAAAGCAATGCAGATAGATGGCGATGGCTCCAGCATC 240
DB 348 CTGAGGGTTCAAAACAAAGGACAGCAT-----GCTCCAGCATC 386

QY 241 AGTTATGAGAAATGGAATGAAGAAGATCCAAAAGTGTCTTTGGGGTGACATAGAGACA 300
DB 387 AGTTATGAGAACTGGTT-----GACCCATTGGAATGTTTATGGTGAGCAGAGACACA 440

QY 301 GGGTTTCATAAGTGGGAGAAATTTTACTGTGAACAACAGATCCTTTTGTCTGCGA---- 356
DB 441 AGGCTTCTGTGAGTGGTTTAAAGTGTGACTGTGAACAACAACTTTCTTATATGCAAGTTC 500

QY 357 --GGCATAGTCTGAAGATCCAGCTGATTAAGTCTGGAAGCAAGAGAGCCGCCACCC 414
DB 501 ACGGACCACTTAAGATCCGGCTGTGTGAGTCTGGAGAGCAAGAGAGCCGCCACCT 560

QY 415 CATCCCCCAACCTCCCTAGACCAATCTCTGCTATGACACCTTTTGTCTCAACGAGATGTC 474
DB 561 C--TCCCCACCCCCACCTTCGCAATCTCTGCTCTTCCCCCTTTGCTCAGTGGATGCTC 618

QY 475 TCTGTAGCTGGATCTGGTGTCTGCTCTCTGATGGCGGAG--TCAATAATCTGCC 532
DB 619 TCTGTAGCGGATCTGGGTTTCTCTCCAGATGGGTGAGAGATCCATAATCTGCC 678

QY 533 TAGCCTGAAAA 544
|||||
```

Db 679 TACCCAAAAAA 690

RESULT 3

US-08-612-840A-6
Sequence 6, Application US/08612840A
Patent No. 5856126
GENERAL INFORMATION:
APPLICANT: FUKUCHI, Naoyuki
APPLICANT: YAMAMOTO, Hitoshi
APPLICANT: NAGANO, Mitsuyo
APPLICANT: KITO, Morikazu
APPLICANT: TANAKA, Akiko
APPLICANT: ISHII, Koichi
APPLICANT: KOBAYASHI, Tsuyoshi
APPLICANT: YOSHIMOTO, Ryota
TITLE OF INVENTION: PEPTIDE HAVING ANTI-THROMBUS ACTIVITY AND
METHOD OF PRODUCING THE SAME
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Obolon, Spivak, McClelland, Maier & Neustadt, P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
City: Arlington
STATE: VA
COUNTRY: USA

ZIP: 22202
 COMPUTER TYPE: FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: FastSeq Version 1.5
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/612,840A
 FILING DATE: 20-MAR-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 5-236975
 FILING DATE: 22-SEP-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 5856126man F. Oblon
 REGISTRATION NUMBER: 24,618
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703-413-3000
 TELEFAX: 703-413-2220
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 272 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: *Crotalus horridus* horrid
 STRAIN:
 US-08-612-840A-6

Query Match	18.9%;	Score 102.6;	DB 2;	Length 272;
Best Local Similarity	72.9%;	Pred. No. 2.6e-23;		
Matches 132;	Conservative 0;	Mismatches 49;	Indels 0;	Gaps 0;
QY	34	AAGACCTGGACAGATGCAGACAGCTTCTGCGAGGACAGAGGTGAACGGGGGGGCATCTGGTC	93	
Db	7	ATGACTTGGGGCGGATGCAGAGAGGTTCTGCTGGAGCAGGCGAAGGGCGGCATCTCTCTC	66	
QY	94	TCTATCGAAAGCTCCGGAGAAGCAGACACTTTGTGGGCCAGTTGATGTCTCAGAAAGATAAG	153	
Db	67	TCTGTGGAACCGCCCTAGAGAGCATCTTTTGGGACATGTGCTCTATGCGAACAAAGAG	126	
QY	154	TCAGCCAAATCCATGTCTGTGATCGGATCGGATCGAGGGCTCAAAACAAAGAAAGCAATGCAGC	213	
Db	127	TACCTTCACAGCTTATCTGGATTGGAATCGAGGGTTCAAAACAAAGGACAGCATGCTCC	186	

Qy	214 A 214
Db	187 A 187

RESULT 4

US-08-232-463-14/c
; Sequence 14, Application US/082322463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Alexandria Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; CLONE: ptzopt-Fls
US-08-232-463-14

	Query Match	10.7%; Score 58.4; DB 1; Length 7218;
	Best Local Similarity	6.2%; Pred. No. 2e-08;
	Matches 26; Conservative	222; Mismatches 168; Indels 0; Gaps 0;
Qy	11	GCTACAAGGTCTTCAACAACTTAAGACCTCGACATGCAGAGAGCTTCGACGAAGC 70
Db	1452	GATAGAAGAATTGGTACRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNR 1393
Qy	71	AGTGTAACGGGGCATCTCGTCTCTATCAAAGCTCCGAGAGACAGACTTTGTGGGCC 130
Db	1392	RRNR 1333
Qy	131	AGTTGATGTCTCAGAGATAAAGTCAGCCAAAATCCATGTCTCGACTCGAGGCTC 190
Db	1332	FRNR 1273
Qy	191	AAAAAAGAAAGCAATGCAGCATAGAGTCGAGCGCATGGCTCCAGCATCAGTTATGAGA 250

[illegible]

RESULT 5
US-09-058-740-12
; Sequence 12, Application US/09058740
; Patent No. 6489451
; GENERAL INFORMATION:
; APPLICANT: Chun Wang, Benjamin XY Li, Xin Cheng,
; Jing Liu, Li-Wen Niu, Wan-Zhi Huang,
; Zhen-Yu Xu, Dan Luo, Lian-Di Kang,
; Jin-Guo Ding, Fang Rong, Yan Liu and
; Hui-Ran Chen
; TITLE OF INVENTION: AN ANTITHROMBOSIS ENZYME FROM THE SNAKE
; VENOM OF AGKISTRODON ACUTUS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 MB

COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Fastseq for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/058,740
FILING DATE: 10-Apr-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Chen, Anthony C.
REGISTRATION NUMBER: 38,673
REFERENCE/DOCKET NUMBER: 233/298
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 454 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 3...440
OTHER INFORMATION: "N" stands for any base.
"Xaa" stands for any amino acid.
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-058-740-12

Query Match	10.7%;	Score 58;	DB 4;	Length 454;
Best Local Similarity	77.8%;	Pred. No. 6.7e-09;		

	Matches	70;	Conservative	0;	Mismatches	20;	Indels	0;	Gaps	0;
Qy	1	GAAGGGCATTGCTACAAAGSTCTTCAACRAATCTTAAGACCTGGACAGATGCAGAGAGCTTC	60							
Db	99	GAAGGGCATTGCTACAAAGCCCTTCGATGAACCTTAAGACCTGGGCAGATGCAGAGAAATTC	158							
Qy	61	TGCACGAAGCAGGTGAACGGGGGGCATCTG	90							
Db	159	TGCACACAAACACAAAGGCAGCCATCTG	188							

```

RESULT 6
US-09-412-554A-3
; Sequence 3, Application US/09412554A
; Patent No. 6355788
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell
; APPLICANT: Ellsworth, Jeff L.
; TITLE OF INVENTION: FOLLISTATIN RELATED PROT
; FILE REFERENCE: 98-50
; CURRENT APPLICATION NUMBER: US/09/412,554A
; CURRENT FILING DATE: 1999-10-05
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 2949
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Degenerate oligonucleotide
; OTHER INFORMATION: zfst2 polypeptide of SEQ
; NAME/KEY: variation
; LOCATION: (1)...(2949)
; OTHER INFORMATION: Each N is independently a
US-09-412-554A-3

```

Query Match	6.4%;	Score 34.6;	DB 4;	Length 2949;
Best Local Similarity	29.8%;	Pred. No. 0.53;		
Matches	68;	Conservative 43;	Mismatches 117;	Indels 0; Gaps 0;
QY	67	AAGCAGGTGAACGGGGGCATCTGGPCTCTATCGAAAGCTCCGGAGAACGACGACITTTGG	126	
Db	292	AARCCNGTWTGYNWSNGAYGGNGARTTYYTAGRAAYCATGTGARTNCAYMGNGN	351	
QY	127	GGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAATCCATGTCCTGGATCGGACTGAGG	186	
Db	352	GCNTGYTTNAARAARCARAARATHACNATHGNTCAAYAYGARGAYTGYTTTYYAARGCN	411	
QY	187	GCTCAAAACAAAGAAAGCAATGCACATAGAGTGGAGCGATGGCTCCAGCATCAGTTAT	246	
Db	412	GAYAARTGYAARACNACGARTAYNSNAARATGAARAYATGTYTNTNGAYTNCARAAY	471	
QY	247	GAGAAATTGGATTGAAGAAGATCCAAAAGTCTCTTGGGGTGCCACATA	294	
Db	472	CAPAAATATATGATGCARAAAYGARAAYCCNAAVGGNGAYGAYATH	519	

RESULT 7
US-08-454-557C-32
; Sequence 32, Application US/08454557C
; Patent No. 5830670
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.

ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,557C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840003
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 590 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-454-557C-32

Query Match 6.3%; Score 34.2; DB 2; Length 590;

Best Local Similarity 52.1%; Pred. No. 0.32;
Matches 190; Conservative 0; Mismatches 148; Indels 27; Gaps 4;

```
QY 8 ATTGCTACAGGCTCTTCAACAATCTAAGACCTGCACAGATGCAGAGAGCTTCTGCACGA 67
   |||||
Db 16 ACTGCTACTCTTTAATGAAGCCGTGAGACCTGGGTTGATGCAGATCTCTATTGC---C 72
   |||||
QY 68 AGCAGTGAAAGGGGGGCGATCTGGTCTCTATGAAAGCTCCGAGAGACGACACTTTGGG 127
   |||||
Db 73 AGAATCATGAATTCGGGCAACCTGGTCTGTGCTCACCAGCCGAGGGTGCCTTTGGG 132
   |||||
QY 128 GCCAGTTGATGCTCAGAGATTAAGTCAAGAAAGCTCCGAGAGACGACACTTTGGG 187
   |||||
Db 133 CCTCACTGAT---TAAGAGAGATGCGACCTGATGACTTCAATGCTGGATGGCCT--- 184
   |||||
QY 188 CTCAAACAAAGAAAGCAATCAGATAGATGAGCGATGGCTCCAGACATCAGTTATG 247
   |||||
Db 185 -CCATGACCCCAAGAAAGCCGCTGGCAGTGGAGCAGTGGTCCCTGTTCTCTACA 243
   |||||
QY 248 AGAATTGGATTGAAGAAATCCAAAA-----AGTGTCTTGGGGTGACA 292
   |||||
Db 244 AGTCTGGGGCATTTGGAGCCCAAGCAGTGTAACTCTGGCTACTGTGTGAGCCTGACCT 303
   |||||
QY 293 TAGAGACAGGTTTTCATAAGTGGGAGAAATTTTACTGTGAACAAAGATCCTTTTGTCT 352
   |||||
Db 304 CAAGCACAGGATTCAGAAATGGAAGGATGTGCCTTGTGAAGACAAAGTTCTCTTTGTCT 363
   |||||
QY 353 GCGAG 357
   |||||
Db 364 GCAAG 368
```

RESULT 8

US-08-340-426D-32
Sequence 32, Application US/08340426D
Patent No. 5948634

GENERAL INFORMATION:

APPLICANT: de la Monte, Suzanne
APPLICANT: Wands, Jack R.
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESS: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.

COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,426D
FILING DATE: 14-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 590 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-340-426D-32

Query Match 6.3%; Score 34.2; DB 2; Length 590;

Best Local Similarity 52.1%; Pred. No. 0.32;
Matches 190; Conservative 0; Mismatches 148; Indels 27; Gaps 4;

```
QY 8 ATTGCTACAGGCTCTTCAACAATCTAAGACCTGCACAGATGCAGAGAGCTTCTGCACGA 67
   |||||
Db 16 ACTGCTACTCTTTAATGAAGCCGTGAGACCTGGGTTGATGCAGATCTCTATTGC---C 72
   |||||
QY 68 AGCAGTGAAAGGGGGGCGATCTGGTCTCTATGAAAGCTCCGAGAGACGACACTTTGGG 127
   |||||
Db 73 AGAATCATGAATTCGGGCAACCTGGTCTGTGCTCACCAGCCGAGGGTGCCTTTGGG 132
   |||||
QY 128 GCCAGTTGATGCTCAGAGATTAAGTCAAGAAAGCTCCGAGAGACGACACTTTGGG 187
   |||||
Db 133 CCTCACTGAT---TAAGAGAGATGCGACCTGATGACTTCAATGCTGGATGGCCT--- 184
   |||||
QY 188 CTCAAACAAAGAAAGCAATCAGATAGATGAGCGATGGCTCCAGACATCAGTTATG 247
   |||||
Db 185 -CCATGACCCCAAGAAAGCCGCTGGCAGTGGAGCAGTGGTCCCTGTTCTCTACA 243
   |||||
QY 248 AGAATTGGATTGAAGAAATCCAAAA-----AGTGTCTTGGGGTGACA 292
   |||||
Db 244 AGTCTGGGGCATTTGGAGCCCAAGCAGTGTAACTCTGGCTACTGTGTGAGCCTGACCT 303
   |||||
QY 293 TAGAGACAGGTTTTCATAAGTGGGAGAAATTTTACTGTGAACAAAGATCCTTTTGTCT 352
   |||||
Db 304 CAAGCACAGGATTCAGAAATGGAAGGATGTGCCTTGTGAAGACAAAGTTCTCTTTGTCT 363
   |||||
QY 353 GCGAG 357
   |||||
Db 364 GCAAG 368
```

RESULT 9

US-08-450-673C-32
Sequence 32, Application US/08450673C
Patent No. 5948888

GENERAL INFORMATION:

APPLICANT: de la Monte, Suzanne
APPLICANT: Wands, Jack R.
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESS: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 500
CITY: Washington

```

; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,673C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 590 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; US-08-450-673C-32

```

```

Query Match      6.3%; Score 34.2; DB 2; Length 590;
Best Local Similarity 52.1%; Pred. No. 0.32;
Matches 190; Conservative 0; Mismatches 148; Indels 27; Gaps 4;

QY 8 ATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCGAGAGCTTCTGCACGA 67
DB 16 ACTGCTACTACTTTAATGAAGACCGTGAGACCTGGGTTGATGCAGATCTCTATTGC---C 72
QY 68 AGCAGGTGAACGGGGGGGATCTGGTCTCTATCGAAAGCTCCGGAGAGCAGACTTTTGTGG 127
DB 73 AGAACATGAATTCCGGGCAACCTGGTGTCTGTGCTCACCAGCCGAGGGTGCCTTTGTGG 132
QY 128 GCCAGTTGATGCTCAGAAAGTAAAGTCAAGCAATCCATGCTCGGAGAGCAGACTTTGTGG 187
DB 133 CCTCACTGAT---TAAGGAGAGTGCACTGATGACTTCAATGCTCGGATTTGSCCT--- 184
QY 188 CTCAAAACAAAGAAAGCAATGCAGATAGAGTGAGCGATGGCTCCAGCATCAGTTATG 247
DB 185 -CCATGACCCCAAAAGAACCGCCGCTGGCACTGGAGCAGTGGGTCCCTGCTCTCTACA 243
QY 248 AGAATTGGAATTGAAGAAGAATCCAAA-----AGTGTCTTGGGGTGACCA 292
DB 244 AGTCCTGGGCAATTGGAGCCCAAGCAGTGTAAATCTCGGCTACTGTGTGAGCCTGACCT 303
QY 293 TAGAGACAGGGTTTCATAAGTGGGAGAAATTTTACTGTGAACAAACAGATCCTTTTGTCT 352
DB 304 CAAGCACAGGATTCAGAAATGGAAGGATGTCCTTTGTGAAGACAAGTTCTCCCTTTGTCT 363
QY 353 GCGAG 357
DB 364 GCAAG 368

```

```

RESULT 10
PCT-US95-17111A-32
; Sequence 32, Application PC/TUS951711A
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and
; TITLE OF INVENTION: Detection of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington

```

```

; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/17111A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/340,426
; FILING DATE: 14-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 590 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; PCT-US95-17111A-32

```

```

Query Match      6.3%; Score 34.2; DB 5; Length 590;
Best Local Similarity 52.1%; Pred. No. 0.32;
Matches 190; Conservative 0; Mismatches 148; Indels 27; Gaps 4;

QY 8 ATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCGAGAGCTTCTGCACGA 67
DB 16 ACTGCTACTACTTTAATGAAGACCGTGAGACCTGGGTTGATGCAGATCTCTATTGC---C 72
QY 68 AGCAGGTGAACGGGGGGGATCTGGTCTCTATCGAAAGCTCCGGAGAGCAGACTTTTGTGG 127
DB 73 AGAACATGAATTCCGGGCAACCTGGTGTCTGTGCTCACCAGCCGAGGGTGCCTTTGTGG 132
QY 128 GCCAGTTGATGCTCAGAAAGTAAAGTCAAGCAATCCATGCTCGGAGAGCAGACTTTGTGG 187
DB 133 CCTCACTGAT---TAAGGAGAGTGCACTGATGACTTCAATGCTCGGATTTGSCCT--- 184
QY 188 CTCAAAACAAAGAAAGCAATGCAGATAGAGTGAGCGATGGCTCCAGCATCAGTTATG 247
DB 185 -CCATGACCCCAAAAGAACCGCCGCTGGCACTGGAGCAGTGGGTCCCTGCTCTCTACA 243
QY 248 AGAATTGGAATTGAAGAAGAATCCAAA-----AGTGTCTTGGGGTGACCA 292
DB 244 AGTCCTGGGCAATTGGAGCCCAAGCAGTGTAAATCTCGGCTACTGTGTGAGCCTGACCT 303
QY 293 TAGAGACAGGGTTTCATAAGTGGGAGAAATTTTACTGTGAACAAACAGATCCTTTTGTCT 352
DB 304 CAAGCACAGGATTCAGAAATGGAAGGATGTCCTTTGTGAAGACAAGTTCTCCCTTTGTCT 363
QY 353 GCGAG 357
DB 364 GCAAG 368

```

```

RESULT 11
US-09-146-969-1
; Sequence 1, Application US/09146969
; Patent No. 6228585
; GENERAL INFORMATION:
; APPLICANT: Dieckgraefe, Brian K.
; TITLE OF INVENTION: Gene Markers for Chronic Mucosal Injury
; FILE REFERENCE: 04255.75314
; CURRENT APPLICATION NUMBER: US/09/146,969
; CURRENT FILING DATE: 1998-09-04

```



```

; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-146-969-1

```

Query Match	6.0%;	Score 32.6;	DB 3;	Length 777;
Best Local Similarity	51.8%;	Pred. No. 1.2;		
Matches 189;	Conservative 0;	Mismatches 149;	Indels 27;	Gaps 4;

QY	8	ATTGCTACAAGGTC	TTCAAACAATCTAAGACCTG	GACAGATGCAGAGAGCTTCTGCACGA	67
Db	219	ACTGCTACTACTT	TAATGAAGACCGTGAGACCT	GGGTGATGCAGATCTCTAATGC---	C 275
QY	68	AGCAGGTGAACGGGGGGC	ATCTGGTCTCTATCGAAAGCT	CCGGAGAAGCAGACCTTTGTGG	127
Db	276	AGAACATGAATTCGGGCA	ACTGGTGTCTGTGCTCACC	AGGCCGAGGGTGCCTTTGTGG	335
QY	128	GCCAGTTGATTGCTCAGA	AGATAAAGTCAGCCAAATCC	ATGCTCGATCGGACTGAGGG	187
Db	336	CCTCACTGAT---	TAAGGAGAGTGGCACTGAT	GACTTCAATGCTCGATTGGCCT----	387
QY	188	CTCAAAACAAGAAAAAG	CAATGCAGCATAGAGTGGAG	CGATGGCTCCAGCATCAGTTATG	247
Db	388	-CCATGACCCCAAAAGA	ACCGCGCTGGCACTGGAG	CAGTGGTCCCTCTCCTACA	446
QY	248	AGAAATTGGATTGAAGA	AGAAATCCAAA-----	AGTGCTTGGGGTGCACA	292
Db	447	AGTCCTTGGGCATTTG	GAGCCCAAGCAGTGTATC	CTGGCTACTGTGTAGCGCTGACCT	506
QY	293	TAGAGACAGGGTTTCA	TAAAGTGGGAGAAATTTT	ACTGTGAACAACAAGATCCCTTTTGTCT	352
Db	507	CAAGCACAGGATCCCA	GAATGGAGGATGTCCCT	TGTGAACAAGATTCTCTCTTTGTAT	566
QY	353	GCGAG	357		
Db	567	GCAAG	571		

RESULT 12
US-08-729-103-2
; Sequence 2, Application US/08729103
; Patent No. 5837841
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN REG PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/729,103
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0138 US
; TELECOMMUNICATION INFORMATION:

```

/ TELEPHONE: 415-855-0555
/
/ TELEFAX: 415-845-4166
/ INFORMATION FOR SEQ ID NO: 2:
/
/ SEQUENCE CHARACTERISTICS:
/
/ LENGTH: 614 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ IMMEDIATE SOURCE:
/
/ LIBRARY: COLNFET02
/ CLONE: J1310334
/ US-08-729-103-2

```

	Query Match	6.0%;	Score 32.4;	DB 2;	Length 614;
	Best Local Similarity	53.1%;	Pred. NO. 1.2;		
	Matches	69;	Conservative 0;	Mismatches 61;	Indels 0; Gaps 0;
QY	8	ATTGCTACAGGCTCTCAACAATCTAAGACCTGGACAGACTGAGAGAGCTTCTGCACGA	67		
Db	254	ATTGCTATGGTTACTTTCGGAAGCTGAGAACTGGTCTGATGCCGAGCTCGAGTGTCACT	313		
QY	68	AGCAGCTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAGCAGACACTTTGTGG	127		
Db	314	CTTACGGAACGGAGGCCACCTGGCATCTATCTCTGAGTTTAAGGAAGCCAGCACCATAG	373		
QY	128	GCCAGTTGAT	137		
Db	374	CACAGTACAT	383		

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

US-08-468-413-1

Query Match	6.0%;	Score 32.4;	DB 2;	Length 1114;
Best Local Similarity	53.1%;	Pred. No. 1.7;		
Matches 69;	Conservative 0;	Mismatches 61;	Indels 0;	Gaps 0;
Qy	8	ATTGCTACAAGGTCTTCAAAACAATCTTAAGACCTGGACAGATGCAGAGAGCTTCTGCACGA	67	
Db	229	ATTGCTATGGTTACTTCAGGAAGCTGAGGAACTCGTCTGATGCCGAGCTGAGTGTCAGT	288	
Qy	68	AGCAGGTGAACGGGGGGGCATCTGCTCTCTATCGAAAGCTCCGGAGAAGACAGACTTCTGTGG	127	
Db	289	CTTACGGAACGGAGCCCACTGGCACTCTATCTGAGTTTAAAGGAGCCACCATAG	348	
Qy	128	GCCAGTTGAT	137	
Db	349	CAGAGTTACAT	358	

RESULT 14
 US-09-162-508-1
 ; Sequence 1, Application US/09162508
 ; Patent No. 6080722
 ; GENERAL INFORMATION:
 ; APPLICANT: LI, ET AL.
 ; TITLE OF INVENTION: Human Colon Specific Gene
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
 ; ADDRESSEE: CECCHI, STEWART & OLSTEIN
 ; STREET: 6 BECKER FARM ROAD
 ; CITY: ROSELAND
 ; STATE: NEW JERSEY
 ; COUNTRY: USA
 ; ZIP: 07068
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5 INCH DISKETTE
 ; COMPUTER: IBM PS/2
 ; OPERATING SYSTEM: MS-DOS
 ; SOFTWARE: WORD PERFECT 5.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/162,508
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/468,413
 ; FILING DATE: 06 JUN 95
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: FERRARO, GREGORY D.
 ; REGISTRATION NUMBER: 36,134
 ; REFERENCE/DOCKET NUMBER: 325800-447
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 201-994-1700
 ; TELEFAX: 201-994-1744
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1114 BASE PAIRS
 ; TYPE: NUCLEIC ACID
 ; STRANDEDNESS: SINGLE
 ; TOPOLOGY: LINEAR
 ; MOLECULE TYPE: CDNA
 US-09-162-508-1

	Query Match	6.0%	Score 32.4	DB 3	Length 1114
	Best Local Similarity	53.1%	0	No. 1.7	
	Matches	69	Conservative	0	Mismatches 61; Indels 0; Gaps 0;
Qy	8	ATTGCTCAAGGTCCTTCAAAACAATCTAAGACCTGCACAGATGCAGAGAGCTTCTGCACGA	67		
Db	229	ATTGCTATGGTTACTTTCAGGAAGCTGAGAGAACTGGTCTGATGCCGAGCTCGAGTGTCAGT	288		
Qy	68	AGCAGGTCGAACGGGGGGCATCTGGTCTCTATCGAAGCTCCGGAGAGACAGACTTTGTGG	127		

Db 289 CTTACGGAACGGAGCCACCTGGCATCTATCCTGAGTTTAAAGGAAGCCAGCACCATAG 348

QY 128 GCCAGTTGAT 137

Db 349 CAGAGTACAT 358

RESULT 15
PCT-US95-07169-1
; Sequence 1, Application PC/TUS9507169
; GENERAL INFORMATION:
; APPLICANT: LI, ET AL.
; TITLE OF INVENTION: Human Colon Specific Gene
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07169
; FILING DATE: 06 JUN 95
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FERRARO, GREGORY D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 325800-389
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1114 BASE PAIRS
; TYPE: NUCLEIC ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: CDNA
; PCT-US95-07169-1

	Query Match	6.0%;	Score 32.4;	DB 5;	Length 1114;
	Best local Similarity	53.1%;	Pred. No. 1.7;		
	Matches	69;	Conservative	0;	Mismatches 61; Indels 0; Gaps 0
QY	8	ATTGCTACAAGGTCCTTCAAACAATCTTAAGACCTGCACAGATGCAGAGAGCTTCTGCACA	67		
Db	229	ATTGCTATGTTACTTTCAGGAAGCTGAGGAACCTGGTCTGATGCCGAGCTCGAGTCTCAGT	288		
QY	68	AGCAGGTGAAACGGGGGGGCATCTGGTCTCTATTCGAAAGCTCCGGAGAAGCAGACTTTGTGG	127		
Db	289	CTTACGGAAACGGAGCCCACTCGCATCTATCTCTGAGTTTAAAGGAAGCCAGCACCATAG	348		
QY	128	GCCAGTTGAT	137		
Db	349	CAGAGTACAT	358		

Search completed: December 8, 2003, 12:41:19
Job time : 81 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2003, 12:02:04 ; Search time 346 Seconds
(without alignments)
5225.543 Million cell updates/sec

Title: US-09-938-114-1
Perfect score: 544
Sequence: 1 GAAGGGCATTGTACAGGT.....ATTCTGCCTAGCTGAAAAA 544

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2201672 seqs, 1661799599 residues

Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2.*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	544	100.0	544	11	US-09-938-114-1
2	306.2	56.3	721	10	US-09-929-230-4
3	306.2	56.3	721	13	US-10-226-420-4
4	288	52.9	725	10	US-09-929-230-10
5	288	52.9	725	13	US-10-226-420-10
6	285.8	52.5	580	10	US-09-929-230-7
7	285.8	52.5	580	13	US-10-226-420-7
8	226.4	41.6	690	10	US-09-969-763-2
9	186.6	34.3	456	10	US-09-929-230-6
10	186.6	34.3	456	13	US-10-226-420-6
11	175	32.2	432	10	US-09-929-230-9
12	175	32.2	432	13	US-10-226-420-9
13	166	30.5	474	10	US-09-929-230-12
14	166	30.5	474	13	US-10-226-420-12
15	82.4	15.1	483	10	US-09-929-230-1

Sequence 1, Appli
Sequence 3, Appli
Sequence 3, Appli
Sequence 30, Appli
Sequence 30, Appli
Sequence 1755, Ap
Sequence 1755, Ap
Sequence 1755, Ap
Sequence 3, Appli
Sequence 179264,
Sequence 179264,
Sequence 1761, Ap
Sequence 1761, Ap
Sequence 1761, Ap
Sequence 2993, Ap
Sequence 3, Appli
Sequence 1786, Ap
Sequence 1786, Ap
Sequence 1786, Ap
Sequence 11, Appl
Sequence 340, Appl
Sequence 24, Appl
Sequence 1778, Ap
Sequence 1778, Ap
Sequence 1778, Ap
Sequence 1756, Ap
Sequence 1756, Ap
Sequence 1756, Ap
Sequence 346, App
Sequence 346, App

16 82.4 15.1 483 13 US-10-226-420-1
17 64.2 11.8 453 10 US-09-929-230-3
18 64.2 11.8 453 13 US-10-226-420-3
19 36.8 6.8 1597 9 US-09-726-643-30
20 36.8 6.8 1597 14 US-10-042-141-30
21 36.4 6.7 498 10 US-09-920-300A-1755
22 36.4 6.7 498 13 US-10-099-926-1755
23 36.4 6.7 498 14 US-10-033-528-1755
24 36 6.6 492 9 US-09-801-438-3
25 35.6 6.5 1223197 13 US-10-027-632-179264
26 35.6 6.5 1223197 14 US-10-027-632-179264
27 34.8 6.4 461 10 US-09-920-300A-1761
28 34.8 6.4 461 13 US-10-099-926-1761
29 34.8 6.4 461 14 US-10-033-528-1761
30 34.8 6.4 509 15 US-10-060-036-2993
31 34.6 6.4 2949 15 US-10-013-136-3
32 34.2 6.3 493 10 US-09-920-300A-1786
33 34.2 6.3 493 13 US-10-099-926-1786
34 34.2 6.3 493 14 US-10-033-528-1786
35 34.2 6.3 800 13 US-09-997-003-11
36 34.2 6.3 843 9 US-09-925-301-340
37 34.2 6.3 843 13 US-09-997-003-34
38 33.8 6.2 585 10 US-09-920-300A-1778
39 33.8 6.2 585 13 US-10-099-926-1778
40 33.8 6.2 585 14 US-10-033-528-1778
41 33.8 6.2 593 10 US-09-920-300A-1756
42 33.8 6.2 593 13 US-10-099-926-1756
43 33.8 6.2 593 14 US-10-033-528-1756
44 33.8 6.2 671 15 US-10-184-644-346
45 33.8 6.2 671 15 US-10-184-644-346

ALIGNMENTS

RESULT 1

US-09-938-114-1
; Sequence 1, Application US/09938114
; Publication No. US20030022350A1
; GENERAL INFORMATION:
; APPLICANT: Chun Wang, Benjamin XY Li, Xin Cheng,
; Jing Liu, Li-Wen Niu, Wan-Zhi Huang,
; Zhen-Yu Xu, Dan Luo, Lian-Di Kang,
; Jin-Guo Ding, Fang Rong, Yan Liu and
; Hui-Ran Chen
; TITLE OF INVENTION: AN ANTITHROMBOTIC ENZYME FROM THE SNAKE.
; VENOM OF AGKISTRODON ACUTUS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/938,114
; FILING DATE: 23-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/058,740
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Chen, Anthony C.
; REGISTRATION NUMBER: 38,673
; REFERENCE/DOCKET NUMBER: 233/298

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 544 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-938-114-1

Query Match      100.0%; Score 544; DB 11; Length 544;
Best Local Similarity 100.0%; Pred. No. 1.2e-175;
Matches 544; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAGGGCATTGCTACAAGGTCTTCAACAATCTAAGACTGACACAGATGCAGAGAGCTTC 60
DB 1 GAAGGGCATTGCTACAAGGTCTTCAACAATCTAAGACTGACACAGATGCAGAGAGCTTC 60
QY 61 TGCACGAAGCAGGTGAACGGGGGCATCTGTCTCTATCGAAAGCTCCGGAGAACGAC 120
DB 61 TGCACGAAGCAGGTGAACGGGGGCATCTGTCTCTATCGAAAGCTCCGGAGAACGAC 120
QY 121 TTTGTGGCCCGATTGATCTCAGAGATAAAGTCAGCCAAATCCATCTCTGGATCGGA 180
DB 121 TTTGTGGCCCGATTGATCTCAGAGATAAAGTCAGCCAAATCCATCTCTGGATCGGA 180
QY 181 CTGAGGGCTCAAAACAAAGAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
DB 181 CTGAGGGCTCAAAACAAAGAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
QY 241 AGTTATGAGAAATGGATTGAAGAAAGATCCAAAAAGTGTCTTGGGGTGACATAGAGACA 300
DB 241 AGTTATGAGAAATGGATTGAAGAAAGATCCAAAAAGTGTCTTGGGGTGACATAGAGACA 300
QY 301 GGGTTTCATAGTGGAGAAATTTTACTGTGACACACAGATCCTTTTGTCTGCGAGGCA 360
DB 301 GGGTTTCATAGTGGAGAAATTTTACTGTGACACACAGATCCTTTTGTCTGCGAGGCA 360
QY 361 TAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAGCAAGGAGCCGCCATCC 420
DB 361 TAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAGCAAGGAGCCGCCATCC 420
QY 421 CCAACCCCTGCTAGCCCAATCTCTGCTATGACACCTTTTGTCTCAACGGATGCTCTGTA 480
DB 421 CCAACCCCTGCTAGCCCAATCTCTGCTATGACACCTTTTGTCTCAACGGATGCTCTGTA 480
QY 481 GCTGATCTGGTGTGCTCTCTGATGGCGGGAAGTCAATAAATCTGCTAGCCTGA 540
DB 481 GCTGATCTGGTGTGCTCTCTGATGGCGGGAAGTCAATAAATCTGCTAGCCTGA 540
QY 541 AAAA 544
DB 541 AAAA 544

RESULT 2
US-09-929-230-4
; Sequence 4, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/09/929,230
; CURRENT FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 721

```

```

; TYPE: DNA
; ORGANISM: Sistrurus miliarius
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (91)...(546)
; NAME/KEY: misc feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Zsnk3
US-09-929-230-4

Query Match      56.3%; Score 306.2; DB 10; Length 721;
Best Local Similarity 78.8%; Pred. No. 5e-94;
Matches 365; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 1 GAAGGGCATTGCTACAAGGTCTTCAACAATCTAAGACTGACACAGATGCAGAGAGCTTC 60
DB 187 GATCAGCATTGCTACAGGGTCTTCAACAATCTAAGACTGACACAGATGCAGAGAGGTTC 246
QY 61 TGCACGAAGCAGGTGAACGGGGGCATCTGTCTCTATCGAAAGCTCCGGAGAACGAC 120
DB 247 TGCTCGAGCAGCGGAGGGCGGCATCTCTCTATCGAAAGCTCCGAGAACGACGCC 306
QY 121 TTTGTGGCCCGATTGATTGCTCAGAGATAAAGTCAGCCAAATCCATCTCTGGATCGGA 180
DB 307 TTTGTGGCCCGATTGATTGCTCAGAGATAAAGTCAGCCAAATCCATCTCTATATCTGGATCGGA 366
QY 181 CTGAGGGCTCAAAACAAAGAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
DB 367 CTGAGGGTTCAGGCGAAAGAGCAATGCAGCGGAAAGTGGAGCGATGGCTCCAGCGTC 426
QY 241 AGTTATGAGAAATGGATTGAAGAAAGATCCAAAAAGTGTCTTGGGGTGACATAGAGACA 300
DB 427 AGTTATGAGAAATGGATTGAAGCAAGATCCAAAAATCTCTTGGGCTGCAACAAGGCACA 486
QY 301 GGGTTTCATAGTGGAGAAATTTTACTGTGACACAAAGATCCTTTTGTCTGCGAGGCA 360
DB 487 AATTATCATAAGTGGTCAATATTTACTGTGAGAAATAAATCCTTTTGTCTGCGAGGCA 546
QY 361 TAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAGCAAGGAGCCGCCATCC 420
DB 547 TAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAGCAAGGAGCCGCCATCC 606
QY 421 CCAACCCCTGCTAGCCCAATCTCTGCTATGACACCTTTTGTCT 463
DB 607 CCCACCTGCTGCATCTGTAGCTGGGATCTGTTCTGTGCTC 649

RESULT 3
US-10-226-420-4
; Sequence 4, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/10/226,420
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 721
; TYPE: DNA
; ORGANISM: Sistrurus miliarius
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (91)...(546)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Zsnk3
US-10-226-420-4

```

Query Match 56.3%; Score 306.2; DB 13; Length 721;
Best Local Similarity 78.8%; Pred. No. 5e-94;
Matches 365; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 1 GAAGGGCATTGCTCAAGGCTTTCAACAACTTAAGACTGACAGATGCAGAGGCTTC 60
DB 187 GATCAGCTATTGCTACAGGGTCTTCAACAACTTAAGACTGACAGATGCAGAGGCTTC 246

QY 61 TGCAGAGCAGGCTGAACCGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAACAGAC 120
DB 247 TGCTCGAGCAGGCGAGGGGGCATCTCGTCTCTATCGAAAGCTCCGGAGAACAGAC 306

QY 121 TTGTGGCCGCTGATGCTCAGAGATAAAGTCAAGCAAAATCCATGCTCGATCGGA 180
DB 307 TTGTGGCCGCTGATGCTCAGAGATAAAGTCAAGCAAAATCCATGCTCGATCGGA 366

QY 181 CTGAGGGCTCAAAACAAAGAAAGCAATGCAGCATAGAGTGGAGGATGGCTCCAGCATC 240
DB 367 CTGAGGGTCAAGGCAGAGAGAGCAATGCAGCGGAAGTGGAGCATGGCTCCAGCGTC 426

QY 241 AGTTATGAGATTGATTAAGAGAAATCCAAAAAGTGTCTTTGGGGTGCACATAGAGACA 300
DB 427 AGTTATGAGAACTGATTAAGAGAAATCCAAAAAGTGTCTTTGGGGTGCACAAAGGCACA 486

QY 301 GGGTTTCATAGTGGAGAAATTTTATCTGTAACAACAAAGATCCTTTTGTCTGCGAGCA 360
DB 487 AATTATCATAGTGGTCAATATTTACTGTGGAGAAATAAATCCTTTTGTCTGCGAGCA 546

QY 361 TAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAAGCAAGGAGCGGCGGCGGCGGCGG 420
DB 547 TAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAAGCAAGGAGCGGCGGCGGCGGCGG 606

QY 421 CCAACCTGCTAGCCCAATCTCTGCTATGACCCCTTTGCTC 463
DB 607 CCAACCTGCTAGCCCAATCTCTGCTATGACCCCTTTGCTC 649

RESULT 4

US-09-929-230-10
; Sequence 10, Application US/0929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/09/929,230
; CURRENT FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 725
; TYPE: DNA
; ORGANISM: Sistrurus miliarius
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (88)...(561)
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Zsnk5

US-09-929-230-10

Query Match 52.9%; Score 288; DB 10; Length 725;
Best Local Similarity 75.9%; Pred. No. 8.8e-88;
Matches 416; Conservative 0; Mismatches 110; Indels 22; Gaps 4;

QY 1 GAAGGGCATTGCTCAAGGCTTTCAACAACTTAAGACTGACAGATGCAGAGGCTTC 60
DB 190 GATCAGTATTGCTACAGGGTCTATCAACGATCAAGACTCGGAGCTGGAGCGGCTTC 249

QY 61 TGCAGAGCAGGCTGAACCGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAACAGAC 120
DB 250 TGCTCGAGCAGGCGAGGGGGCATCTGGGCTCTGTGCAAAACGATGAAGAGCAGTC 309

QY 121 TTGTGGCCGCTGATGCTCAGAGATAAAGTCAAGCAAAATCCATGCTCGATCGGA 180
DB 310 TTGTGGCCGCTGATGCTCAGAGATAAAGTCAAGCAAAATCCATGCTCGATCGGA 369

QY 181 CTGAGGGCTCAAAACAAAGAAAGCAATGCAGCATAGAGTGGAGGATGGCTCCAGCATC 240
DB 370 CTGAGGATTTCAAAACAAAGGACAGCAATGCAGCAGAAAGTGGAGCGATGGCTCCAGCGTC 429

QY 241 AGTTATGAGATTGATTAAGAGAAATCCAAAAAGTGTCTTTGGGGTGCACATAGAGACA 300
DB 430 AGTTATGAGAACTGTTAATCAATTCGAAAGTGTCTTTGGCTGCAAAAGAGACA 489

QY 301 GGGTTTCATAGTGGAGAAATTTTATCTGTAACAACAAAGATCCTTTTGTCTGCGA --- 356
DB 490 GAGTTTCTTCAATGCTACATACCTGCTGCAAGAAAAAACCTTTTCTGCTGCAAGTTC 549

QY 357 --GCATAGTCTGAAGTCCAGCTGATTGAAGTCTGGAGAGCAGGAGCGGCGGCGGCGG 414
DB 550 CCGCAGAGTGTAAAGATCCGGCTGTGTGAAGTCTGGAGAGCAGGAAATCCCCCCCCAC 609

QY 415 CATCCCCCAACCCCTGCTAGCCCAATCTCTGCTATGCAACCT-TTGTCTCAACGATGCT 473
DB 610 CGCC-----TGCCACATCTCTGCTGCACTCTGCTCATTCATGATGCT 656

QY 474 CTCTGATGCTGATCTGGTGTGTGCTCTGCTGATGGGCGGGAAG--TCAATAAATTCGTC 531
DB 657 CTCTGATGCTGATCTGGTGTGTGCTCTGCTGATGGGCGGGAAGTCAATAAATTCGTC 716

QY 532 CTAGCCTG 539
DB 717 CTAGCATG 724

RESULT 5

US-10-226-420-10
; Sequence 10, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/10/226,420
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 725
; TYPE: DNA
; ORGANISM: Sistrurus miliarius
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (88)...(561)
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Zsnk5

US-10-226-420-10

Query Match 52.9%; Score 288; DB 13; Length 725;
Best Local Similarity 75.9%; Pred. No. 8.8e-88;
Matches 416; Conservative 0; Mismatches 110; Indels 22; Gaps 4;

QY 1 GAAGGGCATTGCTCAAGGCTTTCAACAACTTAAGACTGACAGATGCAGAGGCTTC 60
DB 190 GATCAGTATTGCTACAGGGTCTATCAACGATCAAGACTCGGAGCTGGAGCGGCTTC 249

QY 61 TGCAGAGCAGGCTGAACCGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAACAGAC 120
DB 250 TGCTCGAGCAGGCGAGGGGGCATCTGGGCTCTGTGCAAAACGATGAAGAGCAGTC 309

QY 121 TTGTGGCCGCTGATGCTCAGAGATAAAGTCAAGCAAAATCCATGCTCGATCGGA 180

```
Db 310 TTTTGGCCAGTTGGTCTCGCAACATAAAGCAAAACCAATATATGTCTGGATTGGA 369
Qy 181 CTGAGGCTCAAAACAAAGAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
Db 370 CTGAGGATTCAAAACAAAGGACAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCGTC 429
Qy 241 AGTTATGAGATTGATTAAGAGAAATCCAAAGAGTGTCTGGGGTGCACATAGAGACA 300
Db 430 AGTTATGAGAACCTGTTAAATCAATTCACAAAGAGTGTCTGGGGTGCACAAAGAGACA 489
Qy 301 GGGTTTCAATAGTGGGAGAAATTTTACTGTGAACAACAGATCTTTTGTCTGCGA---- 356
Db 490 GAGTTTCTCAATGATGATCAATGACTGCGAGAGAAACCTTTTGTCTGCGAAGTTC 549
Qy 357 --GGCATAGTCTGAAGATCCAGCTGATTAAGTCTGGAGAAAGCAAGAGCCGCCACCC 414
Db 550 CGCCAGAGTGTTAAGATCCGGCTGTGTGAAGTCTGGAGAAAGCAAGAGTCCGCCCCAC 609
Qy 415 CATCCCCCAACCTGCTAGCCACAACTCTGCTATGACACCT--TTGCTCAACGGATGCT 473
Db 610 CGCC-----TGCACAAATCTGCTCTGCACTCTGCACTTCCATGGATGCT 656
Qy 474 CTCTGTAGTGGATCTGGTGTGTTGCTCTGCTCTGCTGATGGCGCGAAG--TCAATAAAATCTGC 531
Db 657 CTCTGTGCTGGATCTGGTCTGCTGCTGCTGCTGATGGCGCGAGAGTCCAAATAAATCTGCG 716
Qy 532 CTAGCCTG 539
Db 717 CTAGCATG 724
```

RESULT 6

```
US-09-929-230-7
; Sequence 7, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/09/929,230
; CURRENT FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 580
; TYPE: DNA
; ORGANISM: Sistrurus miliarius
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(434)
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Zsnk4
US-09-929-230-7
```

```
Query Match 52.5%; Score 285.8; DB 10; Length 580;
Best Local Similarity 75.2%; Pred. No. 4.4e-87;
Matches 413; Conservative 0; Mismatches 97; Indels 39; Gaps 3;
```

```
Qy 1 GAAGGGCATTTGCTACAAGGTCTTCAACAATCTAAGACCTGCAAGATGCAGAGAGCTTC 60
Db 63 GATCAGTATTGCTACAGGGTTCATCAACAACCTCAGGAGCTGGAGCGATGCAGAGAGTTC 122
Qy 61 TGCAGAGAGAGGTGAACGGGGGCGATCTGTTCTCTATCGAAAGCTCCGAGAGAGCAGAC 120
Db 123 TGCTCGGAGAGCGGCGGCGGATCTGCTCTCTATTGAAAGCGAGGAGAGAGCGCC 182
Qy 121 TTTTGGGCCAGTTGATTGTCTCAGAAGATAAAGTCAGCAAAATCCATGTCTGGATCGGA 180
Db 183 TTTTGGGCCAGTGGTCTGAGAACATCAGCAAAACAAATATGATGTCTGGATCGGA 242
```

```
Qy 181 CTGAGGGCTCAAAACAAAGAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
Db 243 CTGAGGATTCAAGGGCAAGAGAAAGCAATGCAGCACAAAGTGGAGCGATGGCTCCAGCGTC 302
Qy 241 AGTTATGAGATTGATTAAGAGAAATCCAAAGAGTGTCTGGGGTGCACATAGAGACA 300
Db 303 AATATGAGAACCTGATTAACATGCGACCAAAAGTGTCTTGGGCTGAAAAGAGACA 362
Qy 301 GGGTTTCAATAGTGGGAGAAATTTTACTGTGAACAACAGATCTTTTGTCTGCGA---- 356
Db 363 GGGTTTGCACGTCGCGCAATGTTTCACTGTACACAACAAATCTTTTCAATGTGCAAGTTC 422
Qy 357 --GGCATAGTCTGAAGATCCAGCTGATTAAGTCTGGAGAAAGCAAGAGCCGCCACCC 414
Db 423 CGCCAGAGTGTTAAGATCCGGCTGTGTGAAGTCTGGAGAAAGCAAGAGTCCAAATAAATCTTCC 482
Qy 415 CATCCCCCAACCTGCTAGCCACAACTCTGCTATGACACCTTTTGTCTCAACGGATGCT 474
Db 483 CACCGC-----CACCTTTGCTCAACGGATGCT 511
Qy 475 TCTGTAGTGGATCTGGTGTGTTGCTGCTCTCTGATGGCGCGAAG--TCAATAAAATCTGCC 532
Db 512 TCTGTAGTGGATCTGGTGTGTTGCTGCTCTCTGATGGCGCGAAGTCCAAATAAATCTTCC 571
Qy 533 TAGCCTGAA 541
Db 572 TAGCCTGGA 580
```

RESULT 7

```
US-10-226-420-7
; Sequence 7, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/10/226,420
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 580
; TYPE: DNA
; ORGANISM: Sistrurus miliarius
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(434)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Zsnk4
US-10-226-420-7
```

```
Query Match 52.5%; Score 285.8; DB 13; Length 580;
Best Local Similarity 75.2%; Pred. No. 4.4e-87;
Matches 413; Conservative 0; Mismatches 97; Indels 39; Gaps 3;
```

```
Qy 1 GAAGGGCATTTGCTACAAGGTCTTCAACAATCTAAGACCTGCAAGATGCAGAGAGCTTC 60
Db 63 GATCAGTATTGCTACAGGGTTCATCAACAACCTCAGGAGCTGGAGCGATGCAGAGAGTTC 122
Qy 61 TGCAGAGAGAGGTGAACGGGGGCGATCTGTTCTCTATCGAAAGCTCCGAGAGAGCAGAC 120
Db 123 TGCTCGGAGAGCGGCGGCGGATCTGCTCTCTATTGAAAGCGAGGAGAGAGCGCC 182
Qy 121 TTTTGGGCCAGTTGATTGTCTCAGAAGATAAAGTCAGCAAAATCCATGTCTGGATCGGA 180
Db 183 TTTTGGGCCAGTGGTCTGAGAACATCAGCAAAACAAATATGATGTCTGGATCGGA 242
Qy 181 CTGAGGGCTCAAAACAAAGAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
```

Db 243 CTGAGGATTCAAGCGAAGAGCAATGCGACCAAGTGGAGCGATGGCTCCAGGCTC 302
 QY 241 AGTTATGAGATTGAATGAGAAGAAATCCAAAAGTCTCTTTGGGGTGCACATAGAGACA 300
 Db 303 AATTATGAGAACTGATTAAACATCGACCAAAAAGTGTTTTGGGCTGAAAAAGAGACA 362
 QY 301 GGGTTTCAATAGTGGGAGAAATTTTACTGTGAACAAACAGATCCTTTTGTCTGGA----- 356
 Db 363 GGGTTTCCAGTGGCGCAATGTTCACTGTACACAAACAAATCTTTTCATGTGCAAGTTC 422
 QY 357 --GGCATAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGAGCCGCCACCC 414
 Db 423 CCGCCAGAGTGTAAAGTCCGGCTGTGTGAAGTCTGGAGAGCAAGAGCCGCCACCC 482
 QY 415 CATCCCCAACCTCGCTAGCCACAATCTGTCTATGACCCCTTTGTCTCAACGGATGCTC 474
 Db 483 CACCGC-----CACCTTTGTCTCAACGGATGCTC 511
 QY 475 TCTGTAGCTGGATCTGTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 532
 Db 512 TCTGTAGCTGGATCTGTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 571
 QY 533 TAGCCTGAA 541
 Db 572 TAGCCTGGA 580

RESULT 8
 US-09-969-763-2
 ; Sequence 2, Application US/09969763
 ; Publication No. US20020198363A1
 ; GENERAL INFORMATION:
 ; APPLICANT: FUKUCHI, NAOKYUKI
 ; APPLICANT: KITO, MORIKAZU
 ; APPLICANT: KAYAHARA, TAKASHI
 ; APPLICANT: FUTAKI, FUMIE
 ; APPLICANT: ISHIKAWA, KOHJI
 ; APPLICANT: SUZUKI, EIICHIRO
 ; APPLICANT: GONDOH, KEIKO
 ; APPLICANT: SHIMBA, NOBUHISA
 ; APPLICANT: YAMADA, NAOKYUKI
 ; TITLE OF INVENTION: PROTEIN HAVING ANTITHROMBOTIC ACTIVITY AND METHOD FOR PRODUCING IT
 ; FILE REFERENCE: 214760US
 ; CURRENT APPLICATION NUMBER: US/09/969,763
 ; CURRENT FILING DATE: 2000-10-25
 ; PRIOR APPLICATION NUMBER: JP 2000-305279
 ; PRIOR FILING DATE: 2000-10-04
 ; NUMBER OF SEQ ID NOS: 50
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 2
 ; LENGTH: 690
 ; TYPE: DNA
 ; ORGANISM: Crotalus harridus
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (66)..(512)
 ; OTHER INFORMATION:
 US-09-969-763-2

Query Match 41.6%; Score 226.4; DB 10; Length 690;
 Best Local Similarity 70.5%; Pred. No. 1.1e-66;
 Matches 389; Conservative 0; Mismatches 126; Indels 37; Gaps 5;

QY 1 GAAGGGCAATTCCTCAAGGCTTCAAAACAATCTTAAGACCTCGACAGATGCGAGAGCTTC 60
 Db 168 GATCGGATTCTCAAGGCCCTTCAACACAGAGATGACCTGGCGGATGCGAGAGGTTTC 227
 QY 61 TGACAGAGAGCTGTAACGGGGGCACTCTCTCTATCGAAAGCTCGGAGAGAGAGAC 120
 Db 228 TGCTCGGAGAGCGGGAAGGGCGGCATCTCTCTCTGTGCGAAACCGCCCTAGAAAGCATCC 287
 QY 121 TTTGTGGGCCAGTTGATTGTCTCAGAAGATAAGTCAGCCAAATCCATGCTCTGGATCGGA 180

Db 288 TTTGTGGCAATGTGCTCTATGCGAAACAAAGAGTACCTCACAGTTATATCTGGATTGGA 347
 QY 181 CTGAGGGCTCAAAACAAAGAAAGCAATGCGAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
 Db 348 CTGAGGGTTCAAAACAAAGGACAGCCAT-----GCTCCAGCATC 386
 QY 241 AGTTATGAGAAATGGATTGAAGAAATCCAAAAGTGTCTTTGGGGTGCACATAGAGACA 300
 Db 387 AGTTATGAGAACTGGT-----GACCAATTTGAATGTTTTTATGTTGTCGAGCAGACACA 440
 QY 301 GGGTTTCAATAGTGGGAGAAATTTTACTGTGAACAAACAGATCCTTTTGTCTGGA----- 356
 Db 441 AGGCTTCTGTAGTGGTTTAAAGTTGACTGTGAACAAACAAATCTTTTCATATGCAAGTTC 500
 QY 357 --GGCATAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGAGCCGCCACCC 414
 Db 501 ACGGACCAAGTAAAGATCCGGCTGTGTGAAGTCTGGAGAAGCAAGAGCCGCCACCT 560
 QY 415 CATCCCCAACCTCGCTAGCCACAATCTGTCTATGACCCCTTTGTCTCAACGGATGCTC 474
 Db 561 C--TCCCCACGCCCCACCTTCGCAATCTCTGCTCTTCCCTTTGCTCAGTGGATGCTC 618
 QY 475 TCTGTAGCTGGATCTGGTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 532
 Db 619 TCTGTAGCCGATCTGGGTTTCTGCTCCAGATGGGTGAGAAGATCCCAATAAATTTCTGCC 678
 QY 533 TAGCCTGAAAAA 544
 Db 679 TACCCAAAAA 690

RESULT 9
 US-09-929-230-6
 ; Sequence 6, Application US/09929230
 ; Patent No. US20020161203A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sheppard, Paul O.
 ; APPLICANT: Bishop, Paul D.
 ; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
 ; FILE REFERENCE: 00-72
 ; CURRENT APPLICATION NUMBER: US/09/929,230
 ; CURRENT FILING DATE: 2001-08-13
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 6
 ; LENGTH: 456
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: This degenerate nucleotide sequence encodes the
 ; OTHER INFORMATION: amino acid sequence of SEQ ID NO:5.
 ; NAME/KEY: misc feature
 ; LOCATION: 6..9..21..24..30..33..36..39..42..48..51..54..57..60..63,
 ; LOCATION: 66..69..78..81..84..90..93..114..117..129..135..147..153,
 ; LOCATION: 162..171..177..180..186..189..192..201..204..213..216..222,
 ; LOCATION: 225..231..234..237..246..249..252..255..261..276..279
 ; OTHER INFORMATION: n = A,T,C or G
 ; NAME/KEY: misc feature
 ; LOCATION: 282..285..291..309..312..327..330..333..336..339..360,
 ; LOCATION: 366..372..378..381..384..393..396..414..429..441..447..456
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-929-230-6

Query Match 34.3%; Score 186.6; DB 10; Length 456;
 Best Local Similarity 47.1%; Pred. No. 4e-53;
 Matches 169; Conservative 86; Mismatches 104; Indels 0; Gaps 0;
 QY 1 GAAGGGCAATTCCTCAAGGCTTCAAAACAATCTTAAGACCTCGACAGATGCGAGAGCTTC 60
 Db 97 GAYCARCACTGTYTAYMGNTTVAARCARYTNAARACNTGGGAYGAYGCGARMGNTTY 156
 QY 61 TGACAGAGAGAGTGAACGGGGGCATCTGCTCTCTATCGAAAGCTCCGGAGAGAGAC 120

```
Db 157 TGYWSNGARCARGCNGCNGCAYTNGTNSNATHGARNWSNGARGCNCN 216
Qy 121 TTTGTGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAATCCATGTCGTGATCGGA 180
Db 217 TTYGTGNCNCARYTNGTNCNGARAAAYMGNMNGCNCNATHYNTATYATHTGGATHGNC 276
Qy 181 CTGAGGCTCAAAACAAGAAAGCAATGTCAGCATAGAGTGAGGAGTGCTCCAGCATC 240
Db 277 YTNMGNTNCARGGNAARGAARCAATGYWSNGCNAARTGWSNGAYGGSNWSNNGTN 336
Qy 241 AGTTATGAGATTGGATTGAAGAGAAATCCAAAAGTCTCTTGGGGTGCACATAGAGACA 300
Db 337 WSNATYGARAAYTGGATHGARGCNGARNWSNARACNTGYTNGGNYTNCARCAGNACN 396
Qy 301 GGGTTTCATAAGTGGGAGAAATTTTACTGTGTAACAACAAGATCCCTTTTGTCTGCGAGGC 359
Db 397 AAYTAYCAYARTGGGTNAAYATHTATYTGYGNGARATHAAYCCNTTGTGTGARGC 455

RESULT 10
US-10-226-420-6
; Sequence 6, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/10/226,420
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 456
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: This degenerate nucleotide sequence encodes the
; OTHER INFORMATION: amino acid sequence of SEQ ID NO:5.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 6, 9, 21, 24, 30, 33, 36, 39, 42, 48, 51, 54, 57, 60, 63,
; LOCATION: 66, 69, 78, 81, 84, 90, 93, 114, 117, 129, 135, 147, 153,
; LOCATION: 162, 171, 177, 180, 186, 189, 192, 201, 204, 213, 216, 222,
; LOCATION: 225, 231, 234, 237, 246, 249, 252, 255, 261, 276, 279
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 282, 285, 291, 309, 312, 321, 327, 330, 333, 336, 339, 360,
; LOCATION: 366, 372, 378, 381, 384, 393, 396, 414, 429, 441, 447, 456
; OTHER INFORMATION: n = A,T,C or G
US-10-226-420-6

Query Match 34.3%; Score 186.6; DB 13; Length 456;
Best Local Similarity 47.1%; Pred. No. 4e-53;
Matches 169; Conservative 86; Mismatches 104; Indels 0; Gaps 0;

Qy 1 GAAGGGATTGCTACAAGGCTTCTCAACAATCTAAGACTGACAGCTGACAGATGACAGAGCTTC 60
Db 97 GAYCARCAYTGYTAYMNGTNTTAYARCAATYNAARCACTGGAAYGAYGCGNGARNGNTTY 156
Qy 61 TGCAGAACAGAGGTGAACGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAACGAC 120
Db 157 TGYWSNGARCARGCNGCNGCAYTNGTNSNATHGARNWSNGARGCNCN 216
Qy 121 TTTGTGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAATCCATGTCGTGATCGGA 180
Db 217 TTYGTGNCNCARYTNGTNCNGARAAAYMGNMNGCNCNATHYNTATYATHTGGATHGNC 276
Qy 181 CTGAGGCTCAAAACAAGAAAGCAATGTCAGCATAGAGTGAGGAGTGCTCCAGCATC 240
Db 277 YTNMGNTNCARGGNAARGAARCAATGYWSNGCNAARTGWSNGAYGGSNWSNNGTN 336

Query Match 34.3%; Score 186.6; DB 13; Length 456;
Best Local Similarity 47.1%; Pred. No. 4e-53;
Matches 169; Conservative 86; Mismatches 104; Indels 0; Gaps 0;

Qy 1 GAAGGGATTGCTACAAGGCTTCTCAACAATCTAAGACTGACAGCTGACAGATGACAGAGCTTC 60
Db 97 GAYCARCAYTGYTAYMNGTNTTAYARCAATYNAARCACTGGAAYGAYGCGNGARNGNTTY 156
Qy 61 TGCAGAACAGAGGTGAACGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAACGAC 120
Db 157 TGYWSNGARCARGCNGCNGCAYTNGTNSNATHGARNWSNGARGCNCN 216
Qy 121 TTTGTGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAATCCATGTCGTGATCGGA 180
Db 217 TTYGTGNCNCARYTNGTNCNGARAAAYMGNMNGCNCNATHYNTATYATHTGGATHGNC 276
Qy 181 CTGAGGCTCAAAACAAGAAAGCAATGTCAGCATAGAGTGAGGAGTGCTCCAGCATC 240
Db 277 YTNMGNTNCARGGNAARGAARCAATGYWSNGCNAARTGWSNGAYGGSNWSNNGTN 336
```

```
Qy 241 AGTTATGAGATTGGATTGAAGAGAAATCCAAAAGTCTCTTGGGGTGCACATAGAGACA 300
Db 337 WSNATYGARAAYTGGATHGARGCNGARNWSNARACNTGYTNGGNYTNCARCAGNACN 396
Qy 301 GGGTTTCATAAGTGGGAGAAATTTTACTGTGTAACAACAAGATCCCTTTTGTCTGCGAGGC 359
Db 397 AAYTAYCAYARTGGGTNAAYATHTATYTGYGNGARATHAAYCCNTTGTGTGARGC 455

RESULT 11
US-09-929-230-9
; Sequence 9, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/09/929,230
; CURRENT FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 432
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: This degenerate nucleotide sequence encodes the
; OTHER INFORMATION: amino acid sequence of SEQ ID NO:8.
; NAME/KEY: misc feature
; LOCATION: 6, 15, 18, 21, 24, 27, 42, 45, 57, 78, 81, 93, 96, 99, 111,
; LOCATION: 117, 126, 135, 141, 144, 150, 153, 156, 165, 171, 177, 180,
; LOCATION: 186, 189, 195, 201, 231, 240, 243, 246, 255, 273, 276,
; LOCATION: 285, 291, 294, 300, 315, 327, 330, 345, 348, 360
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: 363, 369, 372, 378, 384, 393, 405, 423, 426
; OTHER INFORMATION: n = A,T,C or G
US-09-929-230-9

Query Match 32.2%; Score 175; DB 10; Length 432;
Best Local Similarity 45.7%; Pred. No. 3.7e-49;
Matches 163; Conservative 85; Mismatches 109; Indels 0; Gaps 0;

Qy 1 GAAGGGATTGCTACAAGGCTTCTCAACAATCTAAGACTGACAGCTGACAGATGACAGAGCTTC 60
Db 61 GAYCARATGYTAYMNGTNTATHAARCAATYTNMGNACTGGGAYGAYGCGNGARNGNTTY 120
Qy 61 TGCAGAACAGAGGTGAACGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAACGAC 120
Db 121 TGYWSNGARCARGCNGAARGGNGCAYTNGTNSNATHGARNWSNGARGCNGCN 180
Qy 121 TTTGTGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAATCCATGTCGTGATCGGA 180
Db 181 TTYGTGNCNCARYTNGTNCNGARAAAYATHAARCAARAAAYARTAYGAYGTNTGGATHGNC 240
Qy 181 CTGAGGCTCAAAACAAGAAAGCAATGTCAGCATAGAGTGAGGAGTGCTCCAGCATC 240
Db 241 YTNMGNACTHARGGNGARGAARCAATGYWSNACNAARTGWSNGAYGGSNWSNNGTN 300
Qy 241 AGTTATGAGATTGGATTGAAGAGAAATCCAAAAGTCTCTTGGGGTGCACATAGAGACA 300
Db 301 AAYTAYGARAAYTNTATHAARCAVCNACNABARARTGYTYGNYTNAARAGARACN 360
Qy 301 GGGTTTCATAAGTGGGAGAAATTTTACTGTGTAACAACAAGATCCCTTTTGTCTGCGAG 357
Db 361 GGGTTTGMNACNTGGMGNAAYGTNCATGYACNCARCARAAAYTNTTATGTGTGAAR 417

RESULT 12
US-10-226-420-9
; Sequence 9, Application US/10226420
; Publication No. US20030157686A1
```



```

; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/10/226,420
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 432
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: This degenerate nucleotide sequence encodes the
; OTHER INFORMATION: amino acid sequence of SEQ ID NO:8.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 6, 15, 18, 21, 24, 27, 42, 45, 57, 78, 81, 93, 96, 99, 111,
; LOCATION: 117, 126, 135, 141, 144, 150, 153, 156, 165, 171, 177, 180,
; LOCATION: 186, 189, 195, 198, 201, 231, 240, 243, 246, 255, 273, 276,
; LOCATION: 285, 291, 294, 297, 300, 315, 327, 330, 345, 348, 360
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: 363_369, 372, 378, 384, 393, 405, 423, 426
; OTHER INFORMATION: n = A,T,C or G
; US-10-226-420-9

Query Match      32.2%; Score 175; DB 13; Length 432;
Best Local Similarity 45.7%; Pred. No. 3.7e-49;
Matches 163; Conservative 85; Mismatches 109; Indels 0; Gaps 0;

QY 1 GAAGGGCATTGCTCAAGGTTCTTCAAAACAATCTAAGACCTGACAGATGCGAGAGCTTC 60
Db 61 GAYCARTAYTGYTAYMGNGTNTAARCAATCTGCTCTATCGAAGCTCCGGAGAGCAGAC 120
QY 61 TGCACGAGCAGGTGAACGGGGGCATCTGCTCTATCGAAGCTCCGGAGAGCAGAC 120
Db 121 TGYWSNGARCARGCNGARAGGNGNCAYTNGTNSNATHGARWNGAYGGNGARGCNGCN 180
QY 121 TTTGTGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCAAAATCCATGCTCGATCGGA 180
Db 181 TTYTNGCNCARYTNGTNGCNGARAAATYATHARCARAAATYATGAYTNGTGATHGNN 240
QY 181 CTGAGGGCTCAAAACAAGAAAGCAATGACATAGATGCGAGCGATGGCTCCAGCATC 240
Db 241 YTNMGNATHCARAAAYARGGNCARCTGYWSNACNAARTCGWSNGAYGNGWSNNGTN 300
QY 241 AGTTATGAGATTCGATTGAGAGAAATCCAAAGATCTCTGGGTCACATAGAGACA 300
Db 301 AAYTAYGARAAYTNGTNAARCAATGCGNACNAARARTGYTGGNYTNAARAGARACN 360
QY 301 GGGTTTCATAGTGGGAGAAATTTTACTGTGCAACAACAAGATCCTTTGCTGCGAG 357
Db 361 GGNTTYMGACNTGGMGNAAYGTNCAYTGYACNCARCARAAAYTNTTYATGTGYAAR 417

RESULT 13
US-09-929-230-12
; Sequence 12, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/09/929,230
; CURRENT FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 474

```

```

; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: This degenerate nucleotide sequence encodes the
; OTHER INFORMATION: amino acid sequence of SEQ ID NO:11.
; NAME/KEY: misc feature
; LOCATION: 6, 9, 21, 24, 30, 33, 36, 39, 42, 48, 51, 54, 57, 60, 63,
; LOCATION: 66, 69, 84, 87, 90, 99, 120, 123, 132, 135, 141, 153, 159,
; LOCATION: 168, 177, 183, 186, 192, 195, 198, 201, 219, 222, 228, 231,
; LOCATION: 237, 240, 243, 246, 273, 282, 285, 288, 303, 315, 318
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: 327, 333, 336, 339, 342, 345, 357, 360, 366, 372, 387, 390,
; LOCATION: 402, 411, 426, 447, 453, 465, 468
; OTHER INFORMATION: n = A,T,C or G
; US-09-929-230-12

Query Match      30.5%; Score 166; DB 10; Length 474;
Best Local Similarity 43.4%; Pred. No. 4.8e-46;
Matches 155; Conservative 88; Mismatches 114; Indels 0; Gaps 0;

QY 1 CAAGGGCATTGCTCAAGGTTCTTCAAAACAATCTAAGACCTGACAGATGCGAGAGCTTC 60
Db 103 GAYCARTAYTGYTAYMGNGTNTAARCAATCTGCTCTATCGAAGCTCCGGAGAGCAGAC 162
QY 61 TGCACGAGCAGGTGAACGGGGGCATCTGCTCTATCGAAGCTCCGGAGAGCAGAC 120
Db 163 TGYWSNGARCARGCNGARAGGNGNCAYTNGCWNSTNGARAAAYGAYGARGCNGTN 222
QY 121 TTTGTGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCAAAATCCATGCTCGATCGGA 180
Db 223 TTYTNGCNCARYTNGTNGCNGCAAYATHAARCARAAATYATGAYTNGTGATHGNN 282
QY 181 CTGAGGGCTCAAAACAAGAAAGCAATGACATAGATGCGAGCGATGGCTCCAGCATC 240
Db 283 YTNMGNATHCARAAAYARGGNCARCTGYWSNACNAARTCGWSNGAYGNGWSNNGTN 342
QY 241 AGTTATGAGATTCGATTGAGAGAAATCCAAAGATCTCTGGGTCACATAGAGACA 300
Db 343 WSNTAYGARAAYTNGTNAARWSNCAYWSNAARARTGYTGGNYTNAARAGARACN 402
QY 301 GGGTTTCATAGTGGGAGAAATTTTACTGTGCAACAACAAGATCCTTTGCTGCGAG 357
Db 403 GARTTYTNCARTGGTAYTAACNGAYTGYGARGAARAARAYTNTTYTGTGYAAR 459

RESULT 14
US-10-226-420-12
; Sequence 12, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/10/226,420
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 474
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: This degenerate nucleotide sequence encodes the
; OTHER INFORMATION: amino acid sequence of SEQ ID NO:11.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 6, 9, 21, 24, 30, 33, 36, 39, 42, 48, 51, 54, 57, 60, 63,
; LOCATION: 66, 69, 84, 87, 90, 99, 120, 123, 132, 135, 141, 153, 159,
; LOCATION: 168, 177, 183, 186, 192, 195, 198, 201, 219, 222, 228, 231,
; LOCATION: 237, 240, 243, 246, 273, 282, 285, 288, 303, 315, 318
; OTHER INFORMATION: n = A,T,C or G

```

```
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 327, 333, 336, 339, 342, 345, 357, 360, 366, 372, 387, 390,
; LOCATION: 402, 411, 426, 447, 453, 465, 468
; OTHER INFORMATION: n = A,T,C or G
US-018-226-420-12

Query Match      30.5%; Score 166; DB 13; Length 474;
Best Local Similarity 43.4%; Pred. No. 4.8e-46;
Matches 155; Conservative 88; Mismatches 114; Indels 0; Gaps 0;

Qy 1 GAAGGGCATTGCTACAGGCTTCAACAACTTAAGACCTGCACAGATGCAGAGAGCTTC 60
Db 103 GAYCARATYGTATVGMGTNATHARMGNTNARACNTGGGAYGAYGNGARMGNTTY 162

Qy 61 TGCACGAGCAGGTTGAACGGGGGCGATCTGGTCTCTATCGAAAGCTCCGGGAGAGCAGAC 120
Db 163 TGYWSNGARCARGCNAARGGNGNCAYVYNGNWSNGTNGARAAAYGAYGARGGNGTN 222

Qy 121 TTTGTGGCCAGTTGATTGCTCAGAGATAAGTCAGCCCAAAATCCATGCTGGATCGGA 180
Db 223 TTYTNGCNCARYTNGTNGCNGNAAAYATHAARCARAAAYCARTAYATYGTNTGGATHGN 282

Qy 181 CTGAGGGCTCAAAACAAAGAAAGCAATGCAGCATAGAGTCAGCGATGGCTCCAGCATC 240
Db 283 YTNMGNATHCARAAAYARGGNCARCARTGYWSNACNAARTGWSNGAYGWSNWSNGTN 342

Qy 241 AGTTATGAGAAATGGATTGAAGAGAAATCCAAAAGTCTCTGGGGTGCCATAGAGACA 300
Db 343 WSNAYGARAAYTNGTNAARWSNCAYWSNAARAARTGYTYGGNYTNAARAARGARACN 402

Qy 301 GGGTTTCATAAGTCGGGAGAAATTTTACTGTGCAACAAGATCCCTTTTGTCTGGAG 357
Db 403 GARITTYTNCARTGCTGTAAYACNGAYTGYGARGAARAAYTNTTYTGTGTAAR 459
```

```
RESULT 15
US-09-929-230-1
; Sequence 1, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/09/929,230
; CURRENT FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 483
; TYPE: DNA
; ORGANISM: Sistrurus miliarius
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(455)
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Zenk2
US-09-929-230-1
```

```
Query Match      15.1%; Score 82.4; DB 10; Length 483;
Best Local Similarity 53.9%; Pred. No. 2.3e-17;
Matches 220; Conservative 0; Mismatches 176; Indels 12; Gaps 2;

Qy 1 GAAGGGCATTGCTACAGGCTTCAACAACTTAAGACCTGCACAGATGCAGAGAGCTTC 60
Db 87 GATCAGCATTCCTACAGGCTTTCAGTGAACCTCAAAACCTGGGATGATGCAGAGAGTTTC 146

Qy 61 TGCACGAGCAGGTTGAACGGGGGCGATCTGGTCTCTATCGAAAGCTCCGGGAGAGCAGAC 120
Db 147 TGCTACACAGCAGACAGAGAGAGCCGCTGGCCCTCCATCCACAGCATGTAAGAGAGCT 206
```

```
Qy 121 TTTGTGGCCAGTTGATTGCTCAGAGATAAAGTCAGCCAAATCCATGTCCTGGATCGGA 180
Db 207 TTTGTGGGCAAACTGGGCTCCCAAACCTTTGAAATTCACCTTCCA-----TGTGGATCGGA 260

Qy 181 CTGAGGGCTCAAAACAAAGAAAGCAATGCAGCATAGAGTCGAGCGATGGCTCCAGCATC 240
Db 261 CTGAAGATCTATGGAAG-----AATGCAATGGCAGTCGAGCGATGACACCAACTG 314

Qy 241 AGTTATGAGAAATGGATTGAAGAAAGATCCAAAAGTCTTTGGGGTGCCATAGAGACA 300
Db 315 GACTACAAAGCCTGGACTCGAAGACCCCTATTGTACAGTAATGCTAGTCAAGACAGATAGG 374

Qy 301 GGGTTTCATAAGTCGGGAGAAATTTTACTGTGCAACAAGATCCCTTTTGTCTGGAGGCA 360
Db 375 ATCTTTTGGTTCAATAGAGGTTGCGAAAGACTGTATCTTTTGTCTGCAAGTTCCAGGCA 434

Qy 361 TAGTCTGAAAGATCCAGCTGATTGAAGTCTGAGAGCAAGCAAGGAGCCCC 408
Db 435 CGGTCTGGAGATCCGGCTGTGTGGAAGTCTGCAGAAAGCAAAAGAGCCC 482
```

Search completed: December 8, 2003, 13:54:34
Job time : 348 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2003, 10:20:33 ; Search time 2261 Seconds
(without alignments)
5847.694 Million cell updates/sec

Title: US-09-938-114-1
Perfect score: 544
Sequence: 1 GAGGGCATTGCTACAGGT.....ATTCTGCCTAGCTGAAAAA 544

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hcc.*
9: gb_est1.*
10: gb_est2.*
11: gb_hcc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_phg.*
27: em_gss_vri.*
28: gb_gss1.*
29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	302	55.5	461	12	BM401683 PH012R Sn
2	233.8	43.0	418	12	BM401662 JLG2G06F S
3	222.4	40.9	556	12	BM401668 JLG2H08F S
4	205	37.7	618	12	BM401682 PH012F Sn

5	203.8	37.5	501	12	BM401648
6	195.6	36.0	370	12	BM401634
7	186.8	34.3	458	12	BM401641
8	186	34.2	460	12	BM401640
9	180.6	33.2	511	12	BM401631
C 10	172.8	31.8	657	12	BM401418
11	161.2	29.6	406	12	BM401598
12	159.6	29.3	410	12	BM401605
13	156.2	28.7	453	12	BM401653
14	152.2	28.0	440	12	BM401459
15	151.2	27.8	456	12	BM401659
16	149.4	27.5	440	12	BM401637
17	146.8	27.0	654	12	BM401401
C 18	134	24.6	341	12	BM401402
19	129.4	23.8	248	12	BM401663
20	120.6	22.2	405	12	BM401607
21	112.4	20.7	574	12	BM401417
22	111.2	20.4	388	12	BM401625
23	99.8	18.3	401	12	BM401610
24	99.8	18.3	633	12	BM401686
25	99.6	18.3	365	12	BM401665
26	82	15.1	510	12	BM401667
27	77.4	14.2	321	12	BM401611
28	77.4	14.2	378	12	BM401620
29	75	13.8	475	12	BM401621
30	75	13.8	504	12	BM401628
31	75	13.8	512	12	BM401495
32	74	13.6	301	12	BM401654
33	74	13.6	378	12	BM401618
34	71	13.1	391	12	BM401469
35	70.4	12.9	315	12	BM401503
36	67.6	12.4	497	12	BM401552
37	63.4	11.7	374	12	BM401520
38	63	11.6	383	12	BM401463
39	63	11.6	406	12	BM401586
40	63	11.6	414	12	BM401532
41	63	11.5	501	12	BM401527
42	62.6	11.5	407	12	BM401658
43	60.4	11.1	570	12	BM401652
44	58.6	10.8	343	12	BM401656
45	57	10.5	307	12	BM401567

ALIGNMENTS

RESULT 1
BM401683/c
LOCUS PH012R Snake Bothrops insularis library IL2 Bothrops insularis CDNA
DEFINITION 3' similar to Snake venom C-type lectin, mRNA sequence.
ACCESSION BM401683
VERSION BM401683.1 GI:20376311
KEYWORDS EST.
SOURCE Bothrops insularis (island jararaca)
ORGANISM Bothrops insularis
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea; Viperidae; Crotalinae; Bothrops.

REFERENCE 1 (bases 1 to 461)
Junqueira-de-Azevedo, I.L.M. and Ho, P.L.
A survey of gene expression and diversity in the venom glands of the pitviper snake Bothrops insularis through the generation of expressed sequence tags (ESTs)
Gene 299 (1-2), 279-291 (2002)
22347338
PUBMED 12459276

CONTACT: Paulo Lee Ho
Centro de Biotecnologia
Instituto Butantan
Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
Tel: 55 11 37 26 7222 ext. 2083
Fax: 55 11 37 26 1505

Email: hoplee@usp.br
This EST corresponds to cluster BITL09A (see Reference)
Seq primer: M13R.

FEATURES

Location/Qualifiers
1..461
/organism="Bothrops insularis"
/mol_type="mRNA"
/db_xref="taxon:8723"
/tissue_type="venom glands"
/clone_lib="Snake Bothrops insularis library IL2"
/note="Organ: venom glands; Vector: pGEM11zf+; Site 1: Eco RI; Site 2: Not I; Sug of mRNA from Bothrops insularis venom glands were primed with oligo-(dT) and reverse transcribed to cDNA using Superscript Plasmid System for cDNA Synthesis and Cloning (Life Technologies). The cDNAs were selected by size (350-600 pb and up 600 pb) in agarose gel electrophoresis, linked to Eco RI adapters and directionally cloned in pGEM11zf+ vector (Promega). ESTs were generated from random clones and grouped in unique sequences. The putative identification of each EST or cluster was obtained through Blast searches (e-value < e-05)."

BASE COUNT 104 a 125 c 107 g 120 t 5 others
ORIGIN

Query Match 55.5%; Score 302; DB 12; Length 461;
Best Local Similarity 82.2%; Pred. No. 3.2e-71;
Matches 382; Conservative 0; Mismatches 75; Indels 8; Gaps 3;

QY 74 TGAACGGGGGCATCTGTCTCTATCGAAAGCTCCGAGAGCAGACTTTGTGGCCAGT 133
DB 461 TGAATGGCGGCATCTGTGTTCTATCGAAAGCGCGGAGAGCAGCTTTGTGGCCAGC 402
QY 134 TGATTGCTCAAGATAAAGTCAGCCAAATCCATGTCTGATCGATCGATGAGGCTCAAA 193
DB 401 TGGTCGNTGAGAACAAACACACAGTCGCGAATTGAATGTCTGATCGAATGAGGTTCAAG 342
QY 194 ACAAGAAAGCAATCGACGATAGTGGAGCGATGGCTCCAGCATCATGTTATCAGAAAT 253
DB 341 GCAAGAAAGCAATGAGC---TCGTGGAGCGATGGCTCCAGCGTCAGTTATCAGAACT 285
QY 254 GGATTGAAGAAATCAAAAAGTGTCTTTGGGTGCAATAGACAGAGGGTTTCATAGT 313
DB 284 GGATTAAAGCAGAAATTAACAAATGATTTGGTTTGAAAGAGATCGAGGTTTCGTAAGT 225
QY 314 GGGAGAAATTTTACTGTGAACAAACAGATCCCTTTTGTCTGCGAGCATAGTCTGAAGATC 373
DB 224 GGAGCAATTTAACTGTGAACAAACAGAAATTCCTTTNGTCTGGAGGTATAGTCTGAAGATG 165
QY 374 CAGCTGATTGAAGTCTGGAGAGCAAGGAAGCCGCCACCCCATCCCAACCCCTGCTTA 433
DB 164 CAGCTGATTGAGTCTGGAGAGCAAGGAAGCCGCCACCC---CTCCCAACCCCGCT 108
QY 434 GCCAATCTCTGATGCAACCTTTGTCTCAACGGATGCTCTCTGTAGCTGGATCTGGTG 493
DB 107 GCCCAATCTCTGCTGCCCCCTTCGCTGAACGGATGCTCTCTGTAGCTGGATCTGGTT 48
QY 494 TTGCTGCTCTCTGATGGCGGGAAG--TCAATAAATTTGCTCTAGC 536
DB 47 TTGCTGCTCTCTGATGGCGGAGAGGTCCATAAATTTNGCTAGC 3

RESULT 2
BM401662
LOCUS
DEFINITION J12G08F Snake Bothrops insularis library IL3 Bothrops insularis
cDNA 5' similar to Snake venom C-type lectin, mRNA sequence.
ACCESSION BM401662
VERSION BM401662.1
KEYWORDS EST.
SOURCE Bothrops insularis (island jararaca)
ORGANISM Bothrops insularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
Viperidae; Crotalinae; Bothrops.

REFERENCE

AUTHORS
TITLE

1 (Bases 1 to 418)
Junqueira-de-Azevedo, I.L.M. and Ho, P.L.
A survey of gene expression and diversity in the venom glands of the pitviper snake Bothrops insularis through the generation of expressed sequence tags (ESTs)

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Paulo Lee Ho
Centro de Biotecnologia
Instituto Butantan
Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
Tel: 55 11 37 26 7222 ext. 2083
Fax: 55 11 37 26 1505
Email: hoplee@usp.br
This EST corresponds to cluster BITL07A (see Reference)
Seq primer: M13F.

FEATURES

Location/Qualifiers
1..418
/organism="Bothrops insularis"
/mol_type="mRNA"
/db_xref="taxon:8723"
/tissue_type="venom glands"
/clone_lib="Snake Bothrops insularis library IL3"
/note="Organ: venom glands; Vector: pGEM11zf+; Site 1: Eco RI; Site 2: Not I; Sug of mRNA from Bothrops insularis venom glands were primed with oligo-(dT) and reverse transcribed to cDNA using Superscript Plasmid System for cDNA Synthesis and Cloning (Life Technologies). The cDNAs were selected by size (350-600 pb and up 600 pb) in agarose gel electrophoresis, linked to Eco RI adapters and directionally cloned in pGEM11zf+ vector (Promega). ESTs were generated from random clones and grouped in unique sequences. The putative identification of each EST or cluster was obtained through Blast searches (e-value < e-05)."

BASE COUNT 116 a 90 c 122 g 90 t
ORIGIN

Query Match 43.0%; Score 233.8; DB 12; Length 418;
Best Local Similarity 75.7%; Pred. No. 1.1e-52;
Matches 305; Conservative 0; Mismatches 92; Indels 6; Gaps 1;

QY 36 GACCTGACAGATCGACAGAGCTTCTCGAAGCAGGTGAACGGGGGCATCTGGTCTC 95
DB 13 GAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 72
QY 96 TATCGAAGCTCCGAGAGCAGACTTTGTGGCGAGTTGATTTGCTCAGAGATAAGTC 155
DB 73 TATCGAAGCGATGAAGAGCAGACTTTGTGGCGAGCTGGTCTCGAATCATTAGGAA 132
QY 156 AGCCAAAATCATGCTCTGGATCGGAGCTGAGGGCTCAAAAACAAAGAAAGCAATGAGCAT 215
DB 133 ATCCAATATCTGCTGGATCGGAGCTGAGATTTGAACAACAAACAAACAGCAATGAGCTC 192
QY 216 AGAGTGAGCGATGGCTCCAGCATCAGTTATGAGAAATTTGAAGTTGAAGAAATCCAAAA 275
DB 193 GAAGTGAGCGATTTACTCCAGCGTCAGTTATGAGAACTGGTTCCAGAAATGTGAAGAA 252
QY 276 GTGTCTTTGGGTGCACATAGACAGAGTTTCAATAGTGGGAGAAATTTTACTGTGAACA 335
DB 253 GTGTTTTCGCTGGAAGAAAGCAAGGGTTTCGTAAGTGGGTCAATATTGACTGTGAGA 312
QY 336 ACAAGATCTCTTTGTCTGCGAG-----GCATAGTCTTGAAGATCCAGCTGATTGAAGTCT 389
DB 313 AGGAATATCTTCTGTTGCAAGTTTCATACGACCGGTTAAGATCCGCTGAGTGAAGTCT 372
QY 390 GGAGACAGAGAGAGCCCCCCCCCAACCCCAATCCCAACCCCTGCT 432
DB 373 GGAGAGCAAG 415

```

RESULT 3
BM401668
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

FEATURES
source
1..556
/organism="Bothrops insularis"
/mol_type="mRNA"
/db_xref="taxon:8723"
/tissue_type="venom glands"
/clone_lib="Snake Bothrops insularis library IL3"
/notes="Organ: venom glands; Vector: pGEM11zf+; Site 1: Eco
RI; Site 2: Not I; Sug of mRNA from Bothrops insularis
venom glands were primed with oligo-(dT) and reverse
transcribed to cDNA using Superscript plasmid system for
cDNA Synthesis and Cloning (Life Technologies). The cDNAs
were selected by size (350-600 pb and up 600 pb) in
agarose gel electrophoresis, linked to Eco RI adaptors and
directionally cloned in pGEM11zf+ vector (Promega). ESTs
were generated from random clones and grouped in unique
sequences. The putative identification of each EST or
cluster was obtained through Blast searches (e-value <
e-05)."
```

BASE COUNT 146 a 126 c 154 g 129 t 1 others

Query Match 40.9%; Score 222.4; DB 12; Length 556;
Best Local Similarity 74.3%; Pred. No. 1.5e-49;
Matches 295; Conservative 0; Mismatches 97; Indels 5; Gaps 1;

```

QY 36 GACCTGGACAGATCGAGAGCTTCTGCACGAGCAGTGTAACGGGGGCGATCTGGTCTC 95
Db 159 GAGCTGGTATGATGACAGAGAGGTTCTGCTCGAGCAGCGAAGCGCGCCATCTGGTCTC 218
QY 96 TATCGAAAGCTCCGAGAGCAGACTTTGTGGCCAGTTGATTCCTCAGAGATAAGTC 155
Db 219 TATCGAAAGCGATGAAGAGCAGACTTTGTGGCCAGCTGGTCTCGAATAGGAA 278
QY 156 AGCCAAAATCCATGTCTGGATCGGACTGAGGCTCAAAACAAAGAAAAGCAATGAGCAT 215
Db 279 ATCCAAATACTATGTCTGATCGGACTGAGGATTGAACAAACAAACAGCATGAGCTC 338
QY 216 AGATGGAGCGGATGCTCCAGCATCAGTTATGAGAAATTTGAGTGAAGAGAAATCCAAAA 275
Db 339 GAAGTGGAGCGGATTACTCCAGCGTCAGTTATGAGAAACCTGGTTTCGAGGAAATGTGAAAA 398
```

```

QY 276 GTGCTTGGGTGCACATAGACAGAGGGTTTCATAGTGGGAGAAATTTTACTGTGAACA 335
Db 399 GTGTTTTCGGCTGGAAAAAAGCAAGGGTTTCGTAGTGGGTCAATATTTGACTGTGTAGA 458
QY 336 ACAAGATCTTTTGTCTGCGAG-----GCAATAGTCTGAAGATCCAGCTGATTGAAGTCTG 390
Db 459 AGGAATCTTTTGTCTGTCGCAAGTTATAGGACCGCGTTAAGATCCGCTGAGTGAAGTCTG 518
QY 391 GAGAAGCAAGGAAGCCCCCCCCACCCCATCCCCCAACC 427
Db 519 GAGANCAAGGAAGCCCCCCCCCTTACGCCCCCC 555

RESULT 4
BM401682
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

FEATURES
source
1..618
/organism="Bothrops insularis"
/mol_type="mRNA"
/db_xref="taxon:8723"
/tissue_type="venom glands"
/clone_lib="Snake Bothrops insularis library IL2"
/notes="Organ: venom glands; Vector: pGEM11zf+; Site 1: Eco
RI; Site 2: Not I; Sug of mRNA from Bothrops insularis
venom glands were primed with oligo-(dT) and reverse
transcribed to cDNA using Superscript Plasmid System for
cDNA Synthesis and Cloning (Life Technologies). The cDNAs
were selected by size (350-600 pb and up 600 pb) in
agarose gel electrophoresis, linked to Eco RI adaptors and
directionally cloned in pGEM11zf+ vector (Promega). ESTs
were generated from random clones and grouped in unique
sequences. The putative identification of each EST or
cluster was obtained through Blast searches (e-value <
e-05)."
```

BASE COUNT 157 a 124 c 180 g 149 t 8 others

Query Match 37.7%; Score 205; DB 12; Length 618;
Best Local Similarity 77.1%; Pred. No. 8.3e-45;
Matches 297; Conservative 0; Mismatches 81; Indels 7; Gaps 4;

```

QY 1 GAAGGCGATTGCTCAAGAGTCTTCAAAACAATCTAAGACCTCGACAGATGACAGAGCTTC 60
Db 241 GAAGGCGTCTCTACAGGTTCTTCAAGACAGCGAAGAACTGGGACGATGCANAGGTTG 300
```

```

BM401682
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

FEATURES
source
1..618
/organism="Bothrops insularis"
/mol_type="mRNA"
/db_xref="taxon:8723"
/tissue_type="venom glands"
/clone_lib="Snake Bothrops insularis library IL2"
/notes="Organ: venom glands; Vector: pGEM11zf+; Site 1: Eco
RI; Site 2: Not I; Sug of mRNA from Bothrops insularis
venom glands were primed with oligo-(dT) and reverse
transcribed to cDNA using Superscript Plasmid System for
cDNA Synthesis and Cloning (Life Technologies). The cDNAs
were selected by size (350-600 pb and up 600 pb) in
agarose gel electrophoresis, linked to Eco RI adaptors and
directionally cloned in pGEM11zf+ vector (Promega). ESTs
were generated from random clones and grouped in unique
sequences. The putative identification of each EST or
cluster was obtained through Blast searches (e-value <
e-05)."
```

BASE COUNT 157 a 124 c 180 g 149 t 8 others

Query Match 37.7%; Score 205; DB 12; Length 618;
Best Local Similarity 77.1%; Pred. No. 8.3e-45;
Matches 297; Conservative 0; Mismatches 81; Indels 7; Gaps 4;

```

QY 1 GAAGGCGATTGCTCAAGAGTCTTCAAAACAATCTAAGACCTCGACAGATGACAGAGCTTC 60
Db 241 GAAGGCGTCTCTACAGGTTCTTCAAGACAGCGAAGAACTGGGACGATGCANAGGTTG 300
```

QY 61 TGCAGAGAGCTGAACGGGGGCTCTGCTCTCTATCGAAGCTCCGGAGCAGAC 120
 |||||
 Db 301 TGCTCGGAGGAGTGAATGGCGCCATCTGGTCTCTATCGAAGCGCCGGAGACGCC 360
 |||||
 QY 121 TTTGTGGCCAGCTTGAATGCTCAGAAAGATAAAGTCAGCCAAATCCATGCTCTGGATCGGA 180
 |||||
 Db 361 TTTGTGGCCAGCTGCTCGCTCAGAAACAAACACAGTCGCGAATGAATGCTGGATCGGA 420
 |||||
 QY 181 CTGAGGGCTCAAAACAAAGAAAGCAATGCAGCATAGAGTGAGCGATGGCTCCAGATC 240
 |||||
 Db 421 ATGAGGGTTCAAGGCAAAAGAAACAA-----TGCANCTCGTGAGCGATGGCTCCA-CGTC 475
 |||||
 QY 241 AGTTATGAGATGTAATGAGAGCAATCCAAAAGTCTCTGGGGTCACATAGAGACA 300
 |||||
 Db 476 AGTTATGAGAACTGGATTAAGCANAAATTAATAACATGTAATGGGTGAAAAGAAATGCA 535
 |||||
 QY 301 GGGTTTCTAATAGTGGAGAAATTTTACTGTGAAACCAAGATCCCTTTTGTCTGCGAGGCA 360
 |||||
 Db 536 GGGTTTCTAATAGTGGAGCAATCTTAACCTGTGAACAACNA-ATTCCTTTCTCTGCGAGGTA 594
 |||||
 QY 361 TAGTCTGAAGTCCAGCTGATTGAA 385
 |||||
 Db 595 TA-NCTGAAGATGCAGCTGAGTGAA 618
 |||||

RESULT 5
 BM401648
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

BM401648 501 bp mRNA linear EST 01-MAY-2002
 J12803F Snake Bothrops insularis library IL3 Bothrops insularis
 cDNA 5' similar to Snake venom C-type lectin, mRNA sequence.
 BM401648
 BM401648.1 GI:20376276
 EST.
 Bothrops insularis (island jararaca)
 Bothrops insularis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
 Viperidae; Crotalinae; Bothrops.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 COMMENT

Junqueira-de-Azevedo, I.L.M. and Ho, P.L.
 A survey of gene expression and diversity in the venom glands of
 the pitviper snake Bothrops insularis through the generation of
 expressed sequence tags (ESTs)
 Gene 299 (1-2), 279-291 (2002)
 22347338
 12459276
 Contact: Paulo Lee Ho
 Centro de Biotecnologia
 Instituto Butantan
 Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
 Tel: 55 11 37 26 7222 ext. 2083
 Fax: 55 11 37 26 1505
 Email: hoplee@usp.br
 This EST corresponds to cluster BITL07A (see Reference)
 Seq primer: ML3F.

FEATURES
 source

1..501
 Location/Qualifiers
 /organism="Bothrops insularis"
 /mol_type="mRNA"
 /db_xref="taxon:8723"
 /tissue_type="venom glands"
 /clone_lib="Snake Bothrops insularis library IL3"
 /note="Organ: venom glands; Vector: pGEMilzf+; Site 1: Eco
 RI; Site 2: Not I; Sug of mRNA from Bothrops insularis
 venom glands were primed with oligo-(dT) and reverse
 transcribed to cDNA using Superscript Plasmid System for
 cDNA Synthesis and Cloning (Life Technologies). The cDNAs
 were selected by size (350-600 pb and up 600 pb) in
 agarose gel electrophoresis, linked to Eco RI adapters and
 directionally cloned in pGEMilzf+ vector (Promega). ESTs
 were generated from random clones and grouped in unique
 sequences. The putative identification of each EST or
 cluster was obtained through Blast searches (e-value <

BASE COUNT 134 a 102 c 139 g 122 t 4 others
 ORIGIN

Query Match 37.5%; Score 203.8; DB 12; Length 501;
 Best Local Similarity 75.8%; Pred. No. 1.7e-44;
 Matches 250; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 36 GACCTGGACAGATGCAGAGAGCTTCTGCAGAGAGCTGACGAGAGCTGACGAGAGCTGCTGCTC 95
 |||||
 Db 161 GAGCTGGTATGATGCAGAGAGCTTCTGCTCGAGCAGCGAAGGGCGGCCATCTGCTCTC 220
 |||||

QY 96 TATCGAAGCTCCGAGAGAGCAGACTTTGTGGCCAGTTGTTGCTCAGAGATAAAGTC 155
 |||||
 Db 221 TATCGAAGCGATGAAGAAGCAGACTTTGTGGCCAGCTGCTGCTCGAATAGGGAA 280
 |||||

QY 156 AGCCAAAATCCATGCTCTGGATCGGACTGAGGGCTCAAAACAAAGAAAGCAATGCAGCAT 215
 |||||
 Db 281 ATCCAAATACTATGCTCTGGATCGGACTGAGGATGAAAACAAACAGCAATGCAGCTC 340
 |||||

QY 216 AGAGTGGAGCGATGGCTCCAGATCAGTTATGAGAAATGGATTGAAGAAGATCCAAAA 275
 |||||
 Db 341 GAAAGTGGAGCGATTAATCTCCAGCGTCAGTTATGAGAAACCTGGTTCGAGGAAATGTGAAA 400
 |||||

QY 276 GTGCTCTGGGTGCACATAGAGACAGGGCTTTCATAAGTGGGAGAAATTTTACTGTGAACA 335
 |||||
 Db 401 GTGTTTTCGGCTGGAAAAAAGCAAGGGTTTCNTAAGTGGTCAATATGACTGTGTANA 460
 |||||

QY 336 ACAAGATCCTTTTGTCTGCGAGGATAGTC 365
 |||||
 Db 461 AGGAATCCTTTCNTGTGCAATTCATCGAC 490
 |||||

RESULT 6
 BM401634

LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

BM401634 370 bp mRNA linear EST 01-MAY-2002
 J11G02F Snake Bothrops insularis library IL3 Bothrops insularis
 cDNA 5' similar to Snake venom C-type lectin, mRNA sequence.
 BM401634
 BM401634.1 GI:20376262
 EST.

Bothrops insularis (island jararaca)
 Bothrops insularis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
 Viperidae; Crotalinae; Bothrops.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 COMMENT

Junqueira-de-Azevedo, I.L.M. and Ho, P.L.
 A survey of gene expression and diversity in the venom glands of
 the pitviper snake Bothrops insularis through the generation of
 expressed sequence tags (ESTs)
 Gene 299 (1-2), 279-291 (2002)
 22347338
 12459276
 Contact: Paulo Lee Ho
 Centro de Biotecnologia
 Instituto Butantan
 Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
 Tel: 55 11 37 26 7222 ext. 2083
 Fax: 55 11 37 26 1505
 Email: hoplee@usp.br
 This EST corresponds to cluster BITL07A (see Reference)
 Seq primer: ML3F.

FEATURES
 source

1..370
 Location/Qualifiers
 /organism="Bothrops insularis"
 /mol_type="mRNA"
 /db_xref="taxon:8723"
 /tissue_type="venom glands"
 /clone_lib="Snake Bothrops insularis library IL3"
 /note="Organ: venom glands; Vector: pGEMilzf+; Site 1: Eco
 RI; Site 2: Not I; Sug of mRNA from Bothrops insularis
 venom glands were primed with oligo-(dT) and reverse

Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
 Tel: 55 11 37 26 7222 ext. 2083
 Fax: 55 11 37 26 1505
 Email: hoplee@usp.br
 This EST corresponds to cluster BITL13A (see Reference)
 Seq primer: M13F.

FEATURES

Location/Qualifiers

source

1..460

/organism="Bothrops insularis"
 /mol_type="mRNA"
 /db_xref="taxon:8723"
 /tissue_type="venom glands"
 /clone_lib="Snake Bothrops insularis library IL3"
 /note="Organ: venom glands; Vector: pGEM11zf+; Site 1: Eco RI; Site 2: Not I; Sug of mRNA from Bothrops insularis venom glands were primed with oligo-(dT) and reverse transcribed to cDNA using Superscript Plasmid System for cDNA Synthesis and Cloning (Life Technologies). The cDNAs were selected by size (350-600 pb and up 600 pb) in agarose gel electrophoresis, linked to Eco RI adapters and directionally cloned in pGEM11zf+ vector (Promega). ESTs were generated from random clones and grouped in unique sequences. The putative identification of each EST or cluster was obtained through Blast searches (e-value < e-05)."

BASE COUNT 106 a 106 c 145 g 103 t

ORIGIN

Query Match 34.2%; Score 186; DB 12; Length 460;
 Best Local Similarity 77.6%; Pred. No. 1.1e-39;
 Matches 225; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

Qy 1 GAAGGCAATCTCTACAGCTTTCACAACTTAAGACCTGCAGATGCAGAGCTTC 60
 Db |||||
 Qy 170 GAAGGAGATTCTACAGGCTTCACTGAACCGCAAACTCGCGCGATGCAGAGGTT 229
 Db |||||
 Qy 61 TGCACGAGCAGGTGAACGGGGGCATCTGTCTCTATCGAAAGCTCCGAGAGCAGAC 120
 Db |||||
 Qy 230 TGCTCGAGCAGGAGAGGGCGGCATCTGTCTCTATCGAAGGAGGAGGAGCAGAC 289
 Db |||||
 Qy 121 TTTGTGGCCAGTTGATCTCTCAGAAGATAAGTCAAGCCAAATCCATGTCTGGATCGGA 180
 Db |||||
 Qy 290 TTTGTGGCCAGTGTCTCTCAGAGGATGACAGACCCCAATCCATGTCTGGATCGGA 349
 Db |||||
 Qy 181 CTGAGGCTCAAAACAAGAAAGCAATGCAGCATAGATGAGCGATGGTCCAGCATC 240
 Db |||||
 Qy 350 CTGAGGGGTGAACAAAGGACAGCAATGCAGCTGCAATCGAGCGATGGTCTCTGGCTC 409
 Db |||||
 Qy 241 AGTTATGAGATTGGATTGAAGAAGATCCAAAAGTCTTTGGGGTGCA 290
 Db |||||
 Qy 410 TGTATGAGACCTGTTGACGAGAGACGAAATGTTTGTGTGGA 459
 Db |||||

RESULT 9

BM401631

LOCUS

BM401631 JLI131f Snake Bothrops insularis library IL3 Bothrops insularis
 CDNA 5' similar to Snake venom C-type lectin, mRNA sequence.

ACCESSION

BM401631

VERSION

EST.

KEYWORDS

EST.

SOURCE

Bothrops insularis (island jararaca)

ORGANISM

Bothrops insularis

REFERENCE

1 (bases 1 to 511)

AUTHORS

Junqueira-de-Azevedo, I.L.M. and Ho, P.L.

TITLE

A survey of gene expression and diversity in the venom glands of the pitviper snake Bothrops insularis through the generation of expressed sequence tags (ESTs)

JOURNAL

Gene 299 (1-2), 279-291 (2002)

MEDLINE

22347338

PUBMED

COMMENT

12459276
 Contact: Paulo Lee Ho
 Centro de Biotecnologia
 Instituto Butantan
 Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
 Tel: 55 11 37 26 7222 ext. 2083
 Fax: 55 11 37 26 1505
 Email: hoplee@usp.br
 This EST corresponds to cluster BITL07A (see Reference)
 Seq primer: M13F.

FEATURES

source

1..511

/organism="Bothrops insularis"
 /mol_type="mRNA"
 /db_xref="taxon:8723"
 /tissue_type="venom glands"
 /clone_lib="Snake Bothrops insularis library IL3"
 /note="Organ: venom glands; Vector: pGEM11zf+; Site 1: Eco RI; Site 2: Not I; Sug of mRNA from Bothrops insularis venom glands were primed with oligo-(dT) and reverse transcribed to cDNA using Superscript Plasmid System for cDNA Synthesis and Cloning (Life Technologies). The cDNAs were selected by size (350-600 pb and up 600 pb) in agarose gel electrophoresis, linked to Eco RI adapters and directionally cloned in pGEM11zf+ vector (Promega). ESTs were generated from random clones and grouped in unique sequences. The putative identification of each EST or cluster was obtained through Blast searches (e-value < e-05)."

BASE COUNT 135 a 105 c 149 g 119 t 3 others

ORIGIN

Query Match 33.2%; Score 180.6; DB 12; Length 511;
 Best Local Similarity 76.6%; Pred. No. 3.4e-38;
 Matches 219; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

Qy 36 GACCTGCACATGCAGAGCTTCTGCACGAAGCAGGTGAACGGGGGCATCTGGTCTC 95
 Db |||||
 Qy 225 GAGCTGTATGATGCAGAGGTTCTGCTCGAGCAGCGAAGCGGCCATCTGGTCTC 284
 Db |||||
 Qy 96 TATCGAAGCTCCGGAGAGCAGACTTTGTGGGCGAGTTGATTCAGAGATAAAGTC 155
 Db |||||
 Qy 285 TATCGAAGCGATGAAGAAGCAGACTTTGTGGCCANCTGTCGTCGCAATAGGGA 344
 Db |||||
 Qy 156 AGCCAAATCATCTCGATCGGACTCAGGCTCAAAACAAAGAAAGCAATCAGCAT 215
 Db |||||
 Qy 345 ATCCAAATCATCTCGATCGGACTCAGGATGAAACAAACAAACAAACAAATCAGCTC 404
 Db |||||
 Qy 216 AGAGTGGAGCGATGGCTCCAGCATCAGTTATGAGAAATTGAGATGAAGAAGATCAAAA 275
 Db |||||
 Qy 405 GAAGTGGAGCGATTACTCCAGCGTCAGTTATGAGAACCTGTTTCAGGAAATGTGAAAA 464
 Db |||||
 Qy 276 GTGCTTGGGTCACATAGACAGAGGTTTCATAGTGGGAGAT 321
 Db |||||
 Qy 465 GTGTTTTGCGTGGAAAAAAGCAAGGGTTTCNTAAGTGGTCAAT 510
 Db |||||

RESULT 10

BM401418/c

LOCUS

BM401418 GH061R Snake Bothrops insularis library IL2 Bothrops insularis CDNA
 3' similar to Snake venom C-type lectin, mRNA sequence.

ACCESSION

BM401418

VERSION

EST.

KEYWORDS

EST.

SOURCE

Bothrops insularis (island jararaca)

ORGANISM

Bothrops insularis

REFERENCE

1 (bases 1 to 657)

AUTHORS

Junqueira-de-Azevedo, I.L.M. and Ho, P.L.

TITLE

A survey of gene expression and diversity in the venom glands of

the pitviper snake Bothrops insularis through the generation of expressed sequence tags (ESTs)
 Gene 299 (1-2), 279-291 (2002)
 22347338
 12459276
 Contact: Paulo Lee Ho
 Instituto Butantan
 Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
 Tel: 55 11 37 26 7222 ext. 2083
 Fax: 55 11 37 26 1505
 Email: hopleeusp.br
 This EST corresponds to cluster BITL10A (see Reference)
 Seq primer: M13R.

FEATURES

source
 1..857
 /organism="Bothrops insularis"
 /mol_type="mRNA"
 /db_xref="taxon:8723"
 /clone_lib="Snake Bothrops insularis library IL2"
 /note="Organ: venom glands; Vector: pGEM1zf+; Site 1: Eco RI; Site 2: Not I; Sug of mRNA from Bothrops insularis venom glands were primed with oligo-(dT) and reverse transcribed to cDNA using Superscript plasmid System for cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs were selected by size (350-600 pb and up 600 pb) in agarose gel electrophoresis, linked to Eco RI adapters and directionally cloned in pGEM1zf+ vector (Promega). ESTs were generated from random clones and grouped in unique sequences. The putative identification of each EST or cluster was obtained through Blast searches (e-value < e-05)."
 160 a 172 c 152 g 157 t 16 others

BASE COUNT 160 a 172 c 152 g 157 t 16 others
 ORIGIN
 Query Match 31.8%; Score 172.8; DB 12; Length 657;
 Best Local Similarity 64.8%; Pred. No. 4.8e-36;
 Matches 351; Conservative 0; Mismatches 167; Indels 24; Gaps 6;

QY 3 AGGGATTGCTACAGGTCTTCAACAACTTAAGACCTGGACAGATGCAGAGAGCTTCG 62
 DB 526 AAGGAGTGTACAGGGTCTTNGAACAAAGATGAATGAGATGAGGAAATTCG 467
 QY 63 CACGAGCAGGTGAACGGGGGATCTGCTCTATCGAAGTCCGGAGAGCAGACTT 122
 DB 466 CACACACAGCAG-ACAGGGACATNNGTCTCTCCAGAGCAGCGAAGAGCAGATT 408
 QY 123 TGTGGGCCAGTTGATGTGTGTCAGAGAGATAAGTCAAGCCAAATCCATGTCGATCGGACT 182
 DB 407 TGTGGTGTGCTCACCTCAC-----CAATTTGAGAGACAGTTTGTNTGGACCGGACT 354
 QY 183 GAGGGCTCAAAACAAAGAAAGCAATGATGATGAGTGGAGCGATGGCTCCAGCATCAG 242
 DB 353 GAGCGATGTCTGGAAG-----GATGCGAGTTCGAGTGGAGCGATGGCAGCGACCTCAG 300
 QY 243 TTATGCAATTGGATTGATGATGAGAGAGATCCAAAGATGCTTTGGGTGGTGCATAGACAGG 302
 DB 299 NTACAGGACAACTACAGTTTGTGTTTTCAGAAATATGATGTGTGTCATCAAGACAAA 240
 QY 303 GTTTCATAGTGGGAGAAATTTTACTGTGAACAAAGATCTTTTCTGTGCG-----A 356
 DB 239 AAATAACAAATGGAGGATATACCTTCGACAAAGTTGGAATATTTCTGTGGAGTTCCA 180
 QY 357 GGCATAGTCTGAGATCCAGTGAATGAAGTGTGGAGAGCAAGGAAGCCCCCACCACCCA 416
 DB 179 GGCTAGTCTGAGATCCGGTGTGTGAAGTCTGGAGATCAAGGAAGCCCCCACCCTC-- 122
 QY 417 TCCCCCAACCTGCTAGCCACATCTCTGCTATGACCCCTTTGCTCAACGGATGCTCTC 476
 DB 121 -CTCCCCACCCCGTNTGGCGCAINTCTGCTCTTCCCCCTTTTGTGCTGAACGGATGCTCTC 63
 QY 477 TGTAGCTGGATCTGGTGTGCTGCTCTGATGGGCCGGAAG--TCATAAATTTCTGCCTA 534

DB 62 TTAGCTGATCTGGTTTGTCTCTCTGATGGCCAGAGGTCCTCAATAATTCGCTTA 3
 QY 535 GC 536
 DB 2 NC 1

RESULT 11

BM401598
 LOCUS JHAP11F Snake Bothrops insularis library IL3 Bothrops insularis
 DEFINITION CDNA 5' similar to Snake venom C-type lectin, mRNA sequence.
 ACCESSION BM401598
 VERSION
 KEYWORDS BM401598.1 GI:20376226
 SOURCE EST.
 ORGANISM Bothrops insularis (island jararaca)
 Bothrops insularis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Lepidodonta; Squamata; Scleroglossa; Serpentes; Colubroidea;
 Viperidae; Crotalinae; Bothrops.

REFERENCE

1 (bases 1 to 406)
 Junqueira-de-Azevedo, I.L.M. and Ho, P.L.
 A survey of gene expression and diversity in the venom glands of the pitviper snake Bothrops insularis through the generation of expressed sequence tags (ESTs)
 Gene 299 (1-2), 279-291 (2002)
 22347338
 PUBMED 12459276

COMMENT

Contact: Paulo Lee Ho
 Centro de Biotecnologia
 Instituto Butantan
 Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
 Tel: 55 11 37 26 7222 ext. 2083
 Fax: 55 11 37 26 1505
 Email: hopleeusp.br
 This EST corresponds to cluster BITL15A (see Reference)
 Seq primer: M13F.

FEATURES

source
 1..406
 /organism="Bothrops insularis"
 /mol_type="mRNA"
 /db_xref="taxon:8723"
 /clone_lib="Snake Bothrops insularis library IL3"
 /note="Organ: venom glands; Vector: pGEM1zf+; Site 1: Eco RI; Site 2: Not I; Sug of mRNA from Bothrops insularis venom glands were primed with oligo-(dT) and reverse transcribed to cDNA using Superscript plasmid System for cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs were selected by size (350-600 pb and up 600 pb) in agarose gel electrophoresis, linked to Eco RI adapters and directionally cloned in pGEM1zf+ vector (Promega). ESTs were generated from random clones and grouped in unique sequences. The putative identification of each EST or cluster was obtained through Blast searches (e-value < e-05)."
 95 a 91 c 126 g 94 t
 BASE COUNT
 ORIGIN

Query Match 29.6%; Score 161.2; DB 12; Length 406;
 Best Local Similarity 79.9%; Pred. No. 5.9e-33;
 Matches 203; Conservative 0; Mismatches 48; Indels 3; Gaps 1;

QY 1 GAAGGGCATTGCTACAGGTCTTCAACATCTTAAGACCTGGACAGATGCAGAGAGCTTC 60
 DB 156 GAAGGGCGTTCCTACAGGTCTTCAAAACACACCGAAACTGGGACGATGCAGAGAGCTTC 215
 QY 61 TGCACGAAGCAGGTGAACCGGGGGCATCTGCTCTCTATCGAAAGTCCGAGAGCAGAC 120
 DB 216 TGCTCGAGGAGGTGATGGCGCATCTGGTCTCTATCGAAAGCCGCGAGAGCAGAC 275
 QY 121 TTGTGGCGCAGTTGATGTGTCAGAGATAAAGTACGCCAAATCCATGTCTGATCGGA 180

Db 276 TTTGTGCCCCAGCTGGTCACTGAGAACATACAGAGCCGCGAATTTGTCTGATCGGA 335
 QY 181 CTGAGGGCTCAAAACAAAGAAACCAATGACGATAGAGTGGAGCGATGGCTCCAGCATC 240
 Db 336 ATGAGGTTCAAGCAAGAAAGCAATGACG---TCGTGGAGGATGGTCCAGCGTC 392
 QY 241 AGTTATGAGAATG 254
 Db 393 AGTTATGAGAACTG 406

RESULT 12
 BM401605 410 bp mRNA linear EST 01-MAY-2002
 LOCUS JH4H07F Snake Bothrops insularis library IL3 Bothrops insularis
 DEFINITION cDNA 5' similar to Snake venom C-type lectin, mRNA sequence.

ACCESSION BM401605.1 GI:20376233
 VERSION EST.
 KEYWORDS Bothrops insularis (island jararaca)
 SOURCE Bothrops insularis
 ORGANISM Bothrops insularis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
 Viperidae; Crotalinae; Bothrops.

REFERENCE 1 (bases 1 to 410)
 Junqueira-de-Azevedo, I.L.M. and Ho, P.L.
 A survey of gene expression and diversity in the venom glands of
 the pitviper snake Bothrops insularis through the generation of
 expressed sequence tags (ESTs)

JOURNAL Gene 299 (1-2), 279-291 (2002)
 MEDLINE 22347338
 PUBMED 12459276
 COMMENT Contact: Paulo Lee Ho
 Instituto Butantan
 Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
 Tel: 55 11 37 26 7222 ext. 2083
 Fax: 55 11 37 26 1505
 Email: hoples@usp.br
 This EST corresponds to cluster BITL07A (see Reference)
 Seq primer: M13F.

FEATURES
 source
 1..410
 Location/Qualifiers
 /organism="Bothrops insularis"
 /mol_type="mRNA"
 /db_xref="taxon:8723"
 /clone_lib="Snake Bothrops insularis library IL3"
 /note="Organ: venom glands; Vector: pGEM1zf+; Site_1: Eco RI; Site_2: Not I; 5'ug of mRNA from Bothrops insularis venom glands were primed with oligo-(dT) and reverse transcribed to cDNA using Superscript Plasmid System for cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs were selected by size (350-600 pb and up 600 pb) in agarose gel electrophoresis, linked to Eco RI adapters and directionally cloned in pGEM1zf+ vector (Promega). ESTs were generated from random clones and grouped in unique sequences. The putative identification of each EST or cluster was obtained through Blast searches (e-value < e-05)."
 106 a 87 c 120 g 97 t

BASE COUNT 106 a 87 c 120 g 97 t
 ORIGIN
 Query Match 29.3%; Score 159.6; DB 12; Length 410;
 Best Local Similarity 78.0%; Pred. No. 1.6e-32;
 Matches 192; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 36 GACCTGCAGATGACAGAGCTTCTGCAGAGCAGGTGACGGGGGCATCTGCTC 95
 Db 161 GAGCTGGTATGATGACAGAGGTTCTGCTCGGACAGCGAAGCGGGCCATCTGCTC 220
 QY 96 TATCGAAAGCTCCGGAAGACAGACTTTGTGGCCAGTTGATTGCTCAGAGATAAGTC 155

Db 221 TATCGAAGCGATGAAGAGCAGACTTTGTGGCCAGCTGCTCGACATAGGAA 280
 QY 156 AGCCAAATCATGCTCGATCGGACTCGAGGCTCAAAACAAAGAAAGCAATCAGCAT 215
 Db 281 ATCCAAATCATGCTCGATCGGACTCGAGGTTGAAAAACAAAAACAGCAATCAGCTC 340
 QY 216 AGAGTGGAGGATGGCTCCAGCATGATGAGAATTGGATTGAGAAGAAATCCAAAA 275
 Db 341 GAAGTGGAGCGATTACTCCAGCGCTAGTTATGAGAACCTGTTTCAGGAATGTGAAAA 400
 QY 276 GTGTCT 281
 Db 401 GTGTTT 406

RESULT 13
 BM401653

LOCUS JI2D01F Snake Bothrops insularis library IL3 Bothrops insularis
 DEFINITION cDNA 5' similar to Snake venom C-type lectin, mRNA sequence.

ACCESSION BM401653
 VERSION EST.
 KEYWORDS Bothrops insularis (island jararaca)
 SOURCE Bothrops insularis
 ORGANISM Bothrops insularis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
 Viperidae; Crotalinae; Bothrops.

REFERENCE 1 (bases 1 to 453)
 Junqueira-de-Azevedo, I.L.M. and Ho, P.L.
 A survey of gene expression and diversity in the venom glands of
 the pitviper snake Bothrops insularis through the generation of
 expressed sequence tags (ESTs)

JOURNAL Gene 299 (1-2), 279-291 (2002)
 MEDLINE 22347338
 PUBMED 12459276
 COMMENT Contact: Paulo Lee Ho
 Centro de Biotecnologia
 Instituto Butantan
 Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
 Tel: 55 11 37 26 7222 ext. 2083
 Fax: 55 11 37 26 1505
 Email: hoples@usp.br
 This EST corresponds to cluster BITL07A (see Reference)
 Seq primer: M13F.

FEATURES
 source
 1..453
 Location/Qualifiers
 /organism="Bothrops insularis"
 /mol_type="mRNA"
 /db_xref="taxon:8723"
 /clone_lib="Snake Bothrops insularis library IL3"
 /note="Organ: venom glands; Vector: pGEM1zf+; Site_1: Eco RI; Site_2: Not I; 5'ug of mRNA from Bothrops insularis venom glands were primed with oligo-(dT) and reverse transcribed to cDNA using Superscript Plasmid System for cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs were selected by size (350-600 pb and up 600 pb) in agarose gel electrophoresis, linked to Eco RI adapters and directionally cloned in pGEM1zf+ vector (Promega). ESTs were generated from random clones and grouped in unique sequences. The putative identification of each EST or cluster was obtained through Blast searches (e-value < e-05)."
 120 a 96 c 132 g 104 t 1 others

BASE COUNT 120 a 96 c 132 g 104 t 1 others
 ORIGIN
 Query Match 28.7%; Score 156.2; DB 12; Length 453;
 Best Local Similarity 77.7%; Pred. No. 1.4e-31;
 Matches 198; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 36 GACCTGCAGATGACAGAGCTTCTGCAGAGCAGGTGACGGGGGCATCTGCTC 95

Db 212 GAGCTGGTATGACAGAGGTTCTGCTCGAGCAGGCGAAGGGCGGCATCTGGTCTC 271
 QY 96 TATCGAAAGCTCCGGAGACAGACTTTGTGGGCCAGTTGATTTGCTCAGAAGATAAAGTC 155
 Db 272 TATCGAAAGCGATGAAGACAGACTTTGTGGGCCAGCTGGTCTCCNAAATAGGGA 331
 QY 156 AGCCAAATCCATGCTGTGATCGGACTGGAGGCTCAAAACAAAGAAACAATGAGCAT 215
 Db 332 ATCCAAATACTATGCTGTGATCGGACTGGAGGATGAAAAACAAAAACGAATGAGCTC 391
 QY 216 AGAGTGGAGCGATGGTCCAGCATCAGTTATGAGATTGATTTGAGAGATCCAAAA 275
 Db 392 GAAAGTGGAGCGATTACTCCAGCGTCAATGATGAGAACTGGTTCGAGGAATGTGAAAA 451
 QY 276 GT 277
 Db 452 GT 453

RESULT 14
 BM401459
 LOCUS
 DEFINITION JH1D05F Snake Bothrops insularis library IL3 Bothrops insularis
 cDNA 5', similar to Snake venom C-type lectin, mRNA sequence.
 ACCESSION BM401459
 VERSION BM401459.1 GI:20376087
 KEYWORDS EST.
 SOURCE Bothrops insularis (island jararaca)
 ORGANISM Bothrops insularis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
 Viperidae; Crotalinae; Bothrops.
 REFERENCE 1 (bases 1 to 440)
 Junqueira-de-Azevedo, I.L.M. and Ho, P.L.
 A survey of gene expression and diversity in the venom glands of
 the pitviper snake Bothrops insularis through the generation of
 expressed sequence tags (ESTs)
 JOURNAL Gene 299 (1-2), 279-291 (2002)
 MEDLINE 22347338
 PUBMED 12459276
 COMMENT Contact: Paulo Lee Ho
 Centro de Biotecnologia
 Instituto Butantan
 Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
 Tel: 55 11 37 26 7222 ext. 2083
 Fax: 55 11 37 26 1505
 Email: hoplee@usp.br
 This EST corresponds to cluster BITL07A (see Reference)
 Seq primer: M13F.

FEATURES
 source
 1. 440
 /organism="Bothrops insularis"
 /mol_type="mRNA"
 /db_xref="taxon:8723"
 /tissue_type="venom glands"
 /clone_lib="Snake Bothrops insularis library IL3"
 /note="Organ: venom glands; Vector: pGEM11zf+; Site 1: Eco
 RI; Site 2: Not I; Sug of mRNA from Bothrops insularis
 venom glands were primed with oligo-(dT) and reverse
 transcribed to cDNA using Superscript Plasmid System for
 cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs
 were selected by size (350-600 pb and up 600 pb) in
 agarose gel electrophoresis, linked to Eco RI adapters and
 directionally cloned in pGEM11zf+ vector (Promega). ESTs
 were generated from random clones and grouped in unique
 sequences. The putative identification of each EST or
 cluster was obtained through Blast searches (e-value <
 e-05)."

BASE COUNT 113 a 96 c 130 g 101 t
 ORIGIN

Query Match 28.0%; Score 152.2; DB 12; Length 440;

Best Local Similarity 79.0%; Pred. No. 1.7e-30;
 Matches 181; Conservative 0; Mismatches 48; Indels 0; Gaps 0;
 QY 36 GACCTGGACAGATGACAGAGCTTCTGCACGAAGCAGGTGAACGGGGGCATCTGGTCTC 95
 Db 211 GAGCTGGTATGATGACAGAGAGTTCTGCTCGAGCAGGCGAAGGGCGGCATCTGGTCTC 270
 QY 96 TATCGAAAGCTCCGGAGACAGACTTTGTGGGCCAGTTGATTTGCTCAGAAGATAAAGTC 155
 Db 271 TATCGAAAGCGATGAAGACAGACTTTGTGGGCCAGCTGGTCTCCGAAATAGGGA 330
 QY 156 ACCCAAAATCCATGCTGTGATCGGACTGGAGGCTCAAAACAAAGAAACAATGAGCAT 215
 Db 331 ATCCAAATACTATGCTGTGATCGGACTGGAGGATGAAAAACAAAAACGAATGAGCTC 390
 QY 216 AGAGTGGAGCGATGGTCCAGCATCAGTTATGAGATTGATTTGAGAGATCCAAAA 264
 Db 391 GAAAGTGGAGCGATTACTCCAGCGTCAATGATGAGAACTGGTTCGAGGA 439

RESULT 15
 BM401659
 LOCUS
 DEFINITION JL2F01F Snake Bothrops insularis library IL3 Bothrops insularis
 cDNA 5', similar to Snake venom C-type lectin, mRNA sequence.
 ACCESSION BM401659
 VERSION BM401659.1 GI:20376287
 KEYWORDS EST.
 SOURCE Bothrops insularis (island jararaca)
 ORGANISM Bothrops insularis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
 Viperidae; Crotalinae; Bothrops.
 REFERENCE 1 (bases 1 to 456)
 Junqueira-de-Azevedo, I.L.M. and Ho, P.L.
 A survey of gene expression and diversity in the venom glands of
 the pitviper snake Bothrops insularis through the generation of
 expressed sequence tags (ESTs)
 JOURNAL Gene 299 (1-2), 279-291 (2002)
 MEDLINE 22347338
 PUBMED 12459276
 COMMENT Contact: Paulo Lee Ho
 Centro de Biotecnologia
 Instituto Butantan
 Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
 Tel: 55 11 37 26 7222 ext. 2083
 Fax: 55 11 37 26 1505
 Email: hoplee@usp.br
 This EST corresponds to cluster BITL07A (see Reference)
 Seq primer: M13F.

FEATURES
 source
 1. 456
 /organism="Bothrops insularis"
 /mol_type="mRNA"
 /db_xref="taxon:8723"
 /tissue_type="venom glands"
 /clone_lib="Snake Bothrops insularis library IL3"
 /note="Organ: venom glands; Vector: pGEM11zf+; Site 1: Eco
 RI; Site 2: Not I; Sug of mRNA from Bothrops insularis
 venom glands were primed with oligo-(dT) and reverse
 transcribed to cDNA using Superscript Plasmid System for
 cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs
 were selected by size (350-600 pb and up 600 pb) in
 agarose gel electrophoresis, linked to Eco RI adapters and
 directionally cloned in pGEM11zf+ vector (Promega). ESTs
 were generated from random clones and grouped in unique
 sequences. The putative identification of each EST or
 cluster was obtained through Blast searches (e-value <
 e-05)."

BASE COUNT 117 a 100 c 131 g 104 t 4 others
 ORIGIN

Query Match 27.8%; Score 151.2; DB 12; Length 456;

Best Local Similarity 77.9%; Pred. No. 3.2e-30;
 Matches 180; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY	36	GACCTGGACAGATGACAGAGAGCTTCTGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTC	95
Db	226	GAGCTGGTATGATGACAGAGAGCTTCTGCTCGAGCAGCGAAGGGCGGCCATCTGGTCTC	285
QY	96	TATCGAAAGCTCCGGAGAGCAGACTTTGTGGGCCAGTTGATTGCTCAGAGATAAGTC	155
Db	286	TATCNAAAAGCGATGAAGAAGCAGACTTTGTGGCCCACTGCTCGCAACATANGAA	345
QY	156	AGCCAAATCCATGCTGGATCGGACTGAGGCTCAAAACAAAGAAAGCAATGCAGCAT	215
Db	346	ATCCAAATACTATGCTGGATCGGACTGAGGATTGAAAACAAAACAGCAATGCAGCTC	405
QY	216	AGAGTGGAGCGATGGCTCCAGCATCAGTTATGAGAATTGGATTGAAGAAGA	266
Db	406	GAAGTGGAGCGATTACTCCAGGCTCAGTTATGAGAACCCTGTTTCGAGGAAA	456

Search completed: December 8, 2003, 12:39:48
 Job time : 2262 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:45:44 ; Search time 77.4 Seconds
(without alignments)
264.544 Million cell updates/sec

Title: US-09-938-114-2

Perfect score: 715

Sequence: 1 DCSDDWSYEGHCYKVPKQS.....GFHKWENFYCEQDPFVCEA 129

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A Geneseq 19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	715	100.0	129	24	ABU08798
2	477	66.7	152	23	AAE20179
3	475	66.4	152	22	AAW51543
4	460	64.3	154	23	AAO14520
5	434	60.7	144	23	AAE20180
6	424	59.3	131	14	AAE20180
7	422.5	59.1	133	14	AAE20180
8	410.5	57.4	132	23	ABW79130
9	406	56.8	132	14	AAE20180

10	404	56.5	134	14	AAE20180
11	395	55.2	130	14	AAE20180
12	387	54.1	158	23	AAE20180
13	372	52.0	132	16	AAE20180
14	309	43.2	133	14	AAE20180
15	293	41.0	123	14	AAE20180
16	292.5	40.9	126	23	AAE20180
17	290	40.6	146	22	AAE20180
18	289.5	40.5	126	23	AAE20180
19	288.5	40.3	126	23	AAE20180
20	288.5	40.3	126	23	AAE20180
21	288.5	40.3	126	23	AAE20180
22	288.5	40.3	126	23	AAE20180
23	288.5	40.3	126	23	AAE20180
24	287.5	40.2	126	23	AAE20180
25	286.5	40.1	126	16	AAE20180
26	286.5	40.1	126	21	AAE20180
27	286.5	40.1	126	21	AAE20180
28	286.5	40.1	149	16	AAE20180
29	286.5	40.1	149	21	AAE20180
30	286.5	40.1	149	23	AAE20180
31	285.5	39.9	126	23	AAE20180
32	284.5	39.8	126	23	AAE20180
33	284.5	39.8	126	23	AAE20180
34	284.5	39.8	126	23	AAE20180
35	284.5	39.8	126	23	AAE20180
36	284.5	39.8	126	23	AAE20180
37	283.5	39.7	126	23	AAE20180
38	283.5	39.7	126	23	AAE20180
39	282.5	39.5	126	23	AAE20180
40	282.5	39.5	126	23	AAE20180
41	281.5	39.4	126	23	AAE20180
42	280.5	39.2	126	23	AAE20180
43	274.5	38.4	127	13	AAE20180
44	272.5	38.1	145	23	AAE20180
45	267	37.3	127	16	AAE20180

ALIGNMENTS

RESULT 1

ID ABU08798

ABU08798 standard; Protein; 129 AA.

XX ABU08798;

XX AC

XX DT

XX DE

XX XX

XX KW

XX KW

XX KW

XX KW

XX KW

XX KW

XX KW

XX KW

XX OS

XX XX

XX FH

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

Deinagkistrodon acutus antithrombosis enzyme alpha chain.

Antithrombosis; alpha chain; fibrin hydrolysis; blood clot; enzyme; platelet aggregation; vaso-occlusive disorder; thromboembolic disorder; myocardial infarction; restenosis; cancer; neurodegenerative disease; angioathic thrombosis; cerebral thrombosis; thromboangiitis obliterans; ischaemic cerebral vascular disease; unstable angina; acute thrombosis; unstable stenocardia; pulmonary embolism; deep vein thrombosis; oedema; peripheral arterial occlusion; stroke; atherosclerosis; inflammation; thrombosis.

Deinagkistrodon acutus.

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

Sequence of polypeptide
Pigmy rattlesnake
Vipera palestinae
Sequence of polypeptide
M9-G4 Y58A mutant
Snake venom blood
M12-G4 Y63A mutant
M10-G4 K61A mutant
M13-G4 K66A mutant
M15-G4 R100A mutant
M17-G4 R103A mutant
M20-G4 R108A mutant
M8-G4 D54A mutant
Snake venom derive
Snake venom derive
126-mer wild-type
Snake venom antithrombosis
Snake venom derive
149-mer wild-type
M21-G4 D54A, D101N
M11-G4 E62A mutant
M18-G4 R105A mutant
M19-G4 E106A mutant
M23-G4 D54N, D101A
M24-G4 D54A, D101A
M16-G4 D101A mutant
M25-G4 D101A, E106
M14-G4 Y67A mutant
M22-G4 D54A, D101A
M26-G4 D101A, E106
M7-G4 K20A mutant
Sequence of the pl
Korean adder snake
Vipera palestinae

```

PD 03-DEC-2002.
XX
XX
XX 10-APR-1998; 98US-0058740.
XX
XX 10-APR-1997; 97US-043886P.
XX
XX (HEFE-) HEFEI SIU FUNG USTC PHARM CO LTD.
XX
XX Li BX, Cheng X;
XX
XX WPI; 2003-352116/33.
XX
XX N-PSDB; ABX93668.
XX
XX New purified Agkistrodon actus anti-thrombosis enzyme, useful for
XX preventing and treating vaso-occlusive and thromboembolic disorders,
XX including myocardial infarction, restenosis, and cerebral thrombosis and
XX unstable angina -
XX
XX Claim 12; Column 11; 19pp; English.
XX
XX The invention relates to a new Deinagkistrodon actus anti-thrombosis
XX enzyme, where the enzyme hydrolyses fibrin, dissolves blood clots and
XX prevents platelet aggregation. The anti-thrombosis enzyme was
XX administered to rabbits intravenously. Thrombosis was determined before
XX and after the administration at 0.5, 1.0, 2.0 and 3.0 hours by the
XX Chandler method. The enzyme showed anti-thrombosis activity at 0.5 hour
XX following administration at 0.005 micro/kg and this activity was
XX increased significantly at 1.0 hour and at 0.01 micro /kg. The
XX composition is useful for preventing and treating vaso-occlusive and
XX thromboembolic disorders, including myocardial infarction, restenosis,
XX angioathic thrombosis, cerebral thrombosis, ischaemic cerebral vascular
XX diseases, unstable angina, acute thrombosis, unstable stenocardia,
XX thromboangiitis obliterans, pulmonary embolism, deep vein thrombosis,
XX peripheral arterial occlusion, stroke. It is also useful for treating
XX atherosclerosis, oedema and inflammation, cancer and neurodegenerative
XX diseases. The present sequence represents the amino acid sequence of the
XX Deinagkistrodon actus antithrombosis enzyme alpha chain.
XX
XX Sequence 129 AA;
XX
XX Query Match 100.0%; Score 715; DB 24; Length 129;
XX Best Local Similarity 100.0%; Pred. No. 2.8e-68;
XX Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 DCSSDWSYEGHCYKVFQKSTWTDASFCCTKQVNGCHLVSISSGEADFGQLIAQKIK 60
XX |||||
XX 1 DCSSDWSYEGHCYKVFQKSTWTDASFCCTKQVNGCHLVSISSGEADFGQLIAQKIK 60
XX |||||
XX 61 SAKIHVWIGLRAQNKKEKCSIEWSDGSSISKENWIEESKCLGVHIETGFKWENFYCE 120
XX |||||
XX 61 SAKIHVWIGLRAQNKKEKCSIEWSDGSSISKENWIEESKCLGVHIETGFKWENFYCE 120
XX |||||
XX 121 QODPFVCEA 129
XX |||||
XX 121 QODPFVCEA 129
XX |||||
XX
XX RESULT 2
XX AAE20179
XX ID AAE20179 standard; Protein; 152 AA.
XX
XX AAE20179;
XX
XX 18-JUN-2002 (first entry)
XX
XX Pigmy rattlesnake venom gland protein, Zsnk3.
XX
XX Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;
XX platelet aggregation; Zsnk3.
XX
XX Sistrurus miliarius.
XX
XX Key Location/Qualifiers
XX

```

```

Peptide 1..23
Protein /label= Signal_peptide
Disulfide-bond 24..152
Disulfide-bond /label= Mature_Zsnk3_protein
Disulfide-bond 25..36
Disulfide-bond 53..150
Disulfide-bond 125..142
WO200214364-A2.
21-FEB-2002.
13-AUG-2001; 2001WO-US25310.
14-AUG-2000; 2000US-225072P.
14-AUG-2000; 2000US-225087P.
15-AUG-2000; 2000US-225489P.
15-AUG-2000; 2000US-225490P.
20-DEC-2000; 2000US-356997P.
(ZYMO ) ZYMOGENETICS INC.
Sheppard PO, Bishop PD;
WPI; 2002-269180/31.
N-PSDB; AAD32055.
New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which
affect blood coagulation and platelet aggregation system, useful in
therapy and diagnostics, or as tools in the study of genetics or
molecular biology -
Claim 3; Page 2; 79pp; English.
The invention relates to new pigmy rattlesnake (Sistrurus miliarius)
venom gland proteins, which affect blood coagulation and platelet
aggregation system. The polypeptides, which affect blood coagulation and
platelet aggregation system, are useful in therapy and diagnostics. The
polypeptides are also useful as an educational tool in laboratory
practical kits for courses related to genetics and molecular biology,
protein chemistry and antibody production and analysis. The
polynucleotide or polypeptide can be used as standards or as unknowns
for testing purposes. The polypeptides are also useful in identifying
proteins by western blotting, protein purification, determining the
weight of expressed polypeptides as a ratio to total protein expressed,
identifying peptide cleavage sites, coupling amino and carboxyl terminal
tags, mass spectrometry, circular dichroism to determine conformation or
affinity chromatography columns to purify the protein, cloning or
sequencing. The present sequence is Sistrurus miliarius venom gland
protein, Zsnk3.
Sequence 152 AA;
Query Match 66.7%; Score 477; DB 23; Length 152;
Best Local Similarity 64.3%; Pred. No. 7.8e-43;
Matches 83; Conservative 16; Mismatches 30; Indels 0; Gaps 0;
QY 1 DCSSDWSYEGHCYKVFQKSTWTDASFCCTKQVNGCHLVSISSGEADFGQLIAQKIK 60
DB |||||
DB 24 DCPGWSYDQHCYRVFKQKLTWDDAERFCSEAGGHLVSISSGEAAFAVLVPENRR 83
QY 61 SAKIHVWIGLRAQNKKEKCSIEWSDGSSISKENWIEESKCLGVHIETGFKWENFYCE 120
DB |||||
DB 84 RALLYWIGLRAVQGKEKQCSAKWSGDGSSVSNWIEAESKTCGLGQQTQYHKWNIYCG 143
QY 121 QODPFVCEA 129
DB : |||||
DB 144 EINFVCEA 152
RESULT 3
AAM51543
ID AAM51543 standard; Protein; 152 AA.

```

XX AC AAM51543;
 XX 10-JAN-2002 (first entry)
 DT DE Snake venom blood anticoagulant halyxin A chain.
 XX DE Snake; pit viper; venom; halyxin; anticoagulant; thrombogenesis;
 XX KW thrombosis.
 XX OS Agkistrodon halys.
 XX PN KR2001049671-A.
 XX PD 15-JUN-2001.
 XX PF 29-JUN-2000; 2000KR-0036591.
 XX PR 29-JUN-1999; 99KR-0025105.
 XX PA (BIOB-) BIOUD CO LTD.
 XX PI Jang YS, Jung GH, Kim DS, Koo BH, Son YD;
 XX DR WPI; 2001-637330/73.
 XX DR N-PSDB; AAI711876.

XX PT Halyxin as blood anticoagulation protein separated from snake venom -
 XX PS Claim 2; Page 10; 21pp; Korean.
 XX CC The invention relates to halyxin, a novel protein with very strong
 CC blood anticoagulation activity. The protein was separated from snake
 CC venom of Agkistrodon halys brevicaudus (a Korean pit viper) and can
 CC be used in the treatment of thrombogenesis. The present sequence
 CC is the A chain of halyxin.
 XX SQ

Query Match 66.4%; Score 475; DB 22; Length 152;
 Best Local Similarity 63.6%; Pred. No. 1.3e-42;
 Matches 82; Conservative 17; Mismatches 30; Indels 0; Gaps 0;
 QY 1 DCSDDSSYEGHCYKVFQKQKSTWDAESFCTKQVNGHLYSIESSEADFGVQLIAQKIK 60
 Db 24 DCPGWSYEGHCYKVFQKQKSTWDAESFCTKQVNGHLYSIESSEADFGVQLIAQKIK 83
 QY 61 SAKIHWIGLRAQNKKEKQCSIEWSDGSSISKENWIEESKCLGVHETGPHKWFNFCYCE 120
 Db 84 RYGIYIWLRLVRGKQKQSSQWSGSSVSQVNWIEAESKTCGLGLQKETEPRKWFNFCYCE 143
 QY 121 QDPFVCEA 129
 Db 144 ERNPFVCEA 152

RESULT 4
 ID AAO14520 standard; Protein; 154 AA.
 AC AAO14520;
 XX 16-MAY-2002 (first entry)
 DT DE Korean adder snake venom salmorin A chain protein.
 XX KW Korean adder; salmorin protein A chain; snake venom;
 KW fibrinogen clotting inhibition; thrombosis; prothrombin binding;
 XX thrombin binding; blood coagulation.
 OS Agkistrodon halys brevicaudus.
 XX FH Key Location/Qualifiers
 FT Peptide 1.9
 FT /label= Signal_peptide

FT Peptide 1..23
 FT /label= Signal_peptide
 FT Protein 24..154
 FT /note= "Mature salmorin A chain protein"

PN WO200214514-A1.
 XX 21-FEB-2002.
 XX 26-JUL-2001; 2001WO-KR01277.
 XX 27-JUL-2000; 2000KR-0043470.
 XX (BIOB-) BIOUD CO LTD.
 XX Chung K, Kim D, Koh Y;
 XX WPI; 2002-241907/29.
 XX N-PSDB; AAL42015.

XX New salmorin protein derived from venom of Korean adder Agkistrodon
 PT halys brevicaudus, useful for treating thrombosis by repressing
 PT fibrinogen clotting through repression of activation of prothrombin
 PT into thrombin -
 XX Claim 2; Fig 1A; 30pp; English.

XX The invention comprises the nucleotide and protein sequences of a
 CC salmorin protein derived from the venom of Korean adder. Salmorin protein
 CC is composed of an A chain and a B chain, and has inhibitory activity
 CC against fibrinogen clotting. The salmorin protein of the invention is
 CC useful for treating thrombosis, as it represses fibrinogen clotting
 CC potentially by binding to prothrombin and thrombin so as to delay blood
 CC coagulation. The present amino acid sequence represents the Korean adder
 CC salmorin A chain protein.
 XX SQ

Sequence 154 AA;

Query Match 64.3%; Score 460; DB 23; Length 154;
 Best Local Similarity 60.2%; Pred. No. 5.1e-41;
 Matches 77; Conservative 20; Mismatches 31; Indels 0; Gaps 0;
 QY 2 CSDWSYEGHCYKVFQKQKSTWDAESFCTKQVNGHLYSIESSEADFGVQLIAQKIKS 61
 Db 27 CPGWSYEGHCYKVFQKQKSTWDAESFCTKQVNGHLYSIESSEADFGVQLIAQKIKS 86
 QY 62 AKIHWIGLRAQNKKEKQCSIEWSDGSSISKENWIEESKCLGVHETGPHKWFNFCYCEQ 121
 Db 87 SFPHVWIGLRLVRGKQKQSSQWSGSSVSQVNWIEAESKTCGLGLQKETEPRKWFNFCYCEQ 146
 QY 122 QDPFVCEA 129
 Db 147 RNPVCEA 154

RESULT 5
 ID AAE20180 standard; Protein; 144 AA.
 XX AAE20180;
 XX 18-JUN-2002 (first entry)
 DT DE Pigmy rattlesnake venom gland protein, Zsnk4.
 XX KW Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;
 KW platelet aggregation; Zsnk4.
 XX OS Sistrurus miliarius.
 XX FH Key Location/Qualifiers
 FT Peptide 1.9
 FT /label= Signal_peptide

OS Bothrops jararaca.
 PN EP574621-A1.
 XX
 PD 22-DEC-1993.
 XX
 PF 02-JUN-1992; 92EP-0305046.
 XX
 PR 02-JUN-1992; 92EP-0305046.
 XX
 PA (GENZ) GENZYME CORP.
 PA (UYFU-) UNIV FUJITA HEALTH.
 XX
 PI Fujimura Y, Titani K;
 XX
 DR WPI; 1993-407054/51.
 XX
 XX Use of purified botrocetin and opt. ristocetin - for detection of
 PT von Willebrand factor in serum or plasma, for diagnosis of von
 PT Willebrand's disease
 PT
 PS Disclosure; Page 8; 14pp; English.
 XX
 CC Botrocetin, isolated from Bothrops jararaca venom, is a disulfide-
 CC linked heterodimer comprising an alpha subunit (AAR45156) and a beta
 CC subunit (AAR45157). The protein is used to assay von Willebrand
 CC factor in serum or plasma, for diagnosis of von Willebrand disease.
 CC (Updated on 25-MAR-2003 to correct FN field.)
 XX
 SQ Sequence 133 AA;
 Query Match 59.1%; Score 422.5; DB 14; Length 133;
 Best Local Similarity 56.9%; Pred. No. 4.2e-37;
 Matches 74; Conservative 20; Mismatches 35; Indels 1; Gaps 1;
 QY 1 DCSDDWSYEGHCYKVFQKSTWTDABSFCTKQVNGHLSIE-SSGEADFGQLIAQKI 59
 Db 1 DCPSGWSYEGHCYKVFQKSTWTDABSFCTKQVNGHLSIE-SSGEADFGQLIAQKI 60
 QY 60 KSAKIHWIGLRAQNKQCSIEWSDGSSISKENWIEESKCLGVHIETGFHKWENFYC 119
 Db 61 QSSDLYAWIGLRAQNKQCSIEWSDGSSISKENWIEESKCLGVHIETGFHKWENFYC 120
 QY 120 EQQDPFVCEA 129
 Db 121 AQKNPFVCKS 130
 RESULT 8
 ABB79130
 ID ABB79130 standard; protein; 132 AA.
 AC ABB79130;
 XX
 DT 06-AUG-2002 (first entry)
 XX
 DE Ahylysantinfarctase thrombase I subunit A SEQ ID NO:1.
 XX
 KW Ahylysantinfarctase thrombase I; Chinese Agkistrodon acutus;
 KW snake venom; enzyme; Deinagkistrodon acutus.
 XX
 OS Deinagkistrodon acutus.
 XX
 PN CN1332241-A.
 XX
 PD 23-JAN-2002.
 XX
 PF 29-APR-2001; 2001CN-0115567.
 XX
 PR 29-APR-2001; 2001CN-0115567.
 XX
 PA (KUNM-) KUNMING ZOOLOGY INST CHINESE ACAD SCI.
 XX

PI Xiao C;
 XX
 DR WPI; 2002-281832/33.
 XX
 PT Ahylysantinfarctase thrombase and its production process -
 XX
 PS Claim 4; Page 5 (Disclosure); 8pp; Chinese.
 XX
 CC The present sequence represents subunit A of the ahylysantinfarctase
 CC thrombase I isolated from the snake venom of Chinese Agkistrodon acutus
 CC (Deinagkistrodon acutus). The present invention describes an anionic
 CC exchange column chromatographic process used to separate and purify, and
 CC the repurifying in a fast protein purifying work station, where
 CC Ahylysantinfarctase thrombase I of over 97% purity is obtained from
 CC snake venom of Chinese Agkistrodon acutus. The test shows that the
 CC Ahylysantinfarctase thrombase I consists of two subunits designated A
 CC and B (see ABB79130 and ABB79131 respectively).
 XX
 SQ Sequence 132 AA;
 Query Match 57.4%; Score 410.5; DB 23; Length 132;
 Best Local Similarity 58.6%; Pred. No. 7.9e-36;
 Matches 75; Conservative 18; Mismatches 32; Indels 3; Gaps 2;
 QY 1 DCSDDWSYEGHCYKVFQKSTWTDABSFCTKQVNGHLSIESSGEADFGQLIAQKI 60
 Db 3 NCPFGWSAYD-HCY--IKPKNWDABKFCQEQADGGHLVSIKSGERDFLAQLVQSQIE 59
 QY 61 SAKIHWIGLRAQNKQCSIEWSDGSSISKENWIEESKCLGVHIETGFHKWENFYC 120
 Db 60 SIEHVTWGLRLQNKQCSIEWSDGSSISKENWIEESKCLGVHIETGFHKWENFYC 119
 QY 121 QQDPFVCE 128
 Db 120 QLNPFVCK 127
 RESULT 9
 AAR38226
 ID AAR38226 standard; protein; 132 AA.
 AC AAR38226;
 XX
 DT 25-MAR-2003 (updated)
 DT 01-OCT-1993 (first entry)
 XX
 DE Sequence of polypeptide chain of alboaagregin B1 (AL-B1).
 XX
 KW Venom; snake; platelet-binding protein.
 XX
 OS Trimeresurus albolabris.
 XX
 PN WO9311151-A1.
 XX
 PD 10-JUN-1993.
 XX
 PF 01-DEC-1992; 92WO-US10344.
 XX
 PR 03-DEC-1991; 91US-0803630.
 PR 05-JUN-1992; 92US-0893929.
 XX
 XX (UTEM) UNIV TEMPLE.
 PA
 XX Kirby EP, Peng M;
 PI
 DR WPI; 1993-196991/24.
 XX
 XX New platelet-binding proteins obtained from snake venom - inhibit
 PT binding of von Willebrand factor to platelet membrane
 PT glycoprotein IB, useful in therapy, diagnosis and surgery
 XX
 PS Claim 6; Page 47; 74pp; English.
 XX

CC Alboaggregin B1 and B2 (collectively "AL-B"), may be obtained in
 CC substantially pure chemical form from venom. Each of AL-B1 and -B2
 CC comprise an about 23 kDa protein formed by an about 17 kDa
 CC polypeptide chain crosslinked to an about 14 kDa polypeptide chain
 CC by one or more interchain disulfide bonds. Each of the two
 CC polypeptide chains of AL-B1 and -B2 show strong homology to two of
 CC the chains of AL-A. AL-B1 and -B2 each comprise two non-identical
 CC polypeptide chains.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX SQ Sequence 132 AA;
 Query Match 56.8%; Score 406; DB 14; Length 132;
 Best Local Similarity 54.7%; Pred. No. 2.4e-35;
 Matches 70; Conservative 23; Mismatches 35; Indels 0; Gaps 0;
 QY 1 DCSSDWSYEGHCYKVFQKSTWTDASFCYKQVNGGHLVSISSGSEADFGVQLIAQKIK 60
 Db 1 DCPDSWSSFKQCYQIVKELKTWEDAEXFCSEQANDGHLVSISSYREAVFVAELLSENVK 60
 QY 61 SAKIHWIGLRAQNKQCSIEWSDGSSISKENWIEESKCLGVHITGPHKWFYCE 120
 Db 61 TTKYNWIGLVQNKQCSIEWSDGSSVXYENLIKPNPKKCFVLKKESEFRTWSNVYCE 120
 QY 121 QODPFVCE 128
 Db 121 QKHIFMCK 128
 RESULT 10
 AAR38223
 ID AAR38223 standard; protein; 134 AA.
 XX AC AAR38223;
 XX 25-MAR-2003 (updated)
 DT 01-OCT-1993 (first entry)
 XX Sequence of polypeptide chain of alboaggregin A (AL-A).
 DE Venom; snake; platelet-binding protein.
 XX Trimeresurus albolabris.
 XX WO9311151-Al.
 XX 10-JUN-1993.
 XX 01-DEC-1992; 92WO-US10344.
 XX 03-DEC-1991; 91US-0803630.
 PR 05-JUN-1992; 92US-0893929.
 XX (UTEM) UNIV TEMPLE.
 PA Kirby EP, Peng M;
 PI WPI; 1993-196991/24.
 XX New platelet-binding proteins obtained from snake venom - inhibit
 PT binding of von Willebrand factor to platelet membrane
 PT glycoprotein IB, useful in therapy, diagnosis and surgery
 XX Claim 4; Page 46; 74pp; English.
 XX AL-A may be obtd. from snake venom. It has a mol. wt. of about 45
 CC kDa. AL-A contains two types of polypeptide chains, with mol. wt. of
 CC about 18 kDa and about 15 kDa, respectively. Each of these two types
 CC of chains is actually composed of two subtypes. Thus, AL-A comprises
 CC four non-identical polypeptide chains. A possible variation of
 CC AAR38224 has been found, characterised by Asp at posn. three in lieu
 CC of Cys.
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 134 AA;
 Query Match 56.8%; Score 404; DB 14; Length 134;
 Best Local Similarity 55.9%; Pred. No. 3.9e-35;
 Matches 71; Conservative 22; Mismatches 34; Indels 0; Gaps 0;
 QY 2 CSSDWSYEGHCYKVFQKSTWTDASFCYKQVNGGHLVSISSGSEADFGVQLIAQKIKS 61
 Db 4 CLPWSAYDQCYKVFNEPKNWDARFCAQKADSGHLVSIETWGEADFVAQLISENIQS 63
 QY 62 AKIHWIGLRAQNKQCSIEWSDGSSISKENWIEESKCLGVHITGPHKWFYCEQ 121
 Db 64 KEHVWIGLVQNKQCSIEWSDGSSVXYENLIKPNPKKCFVLKKESEFRTWSNVYCE 123
 QY 122 QODPFVCE 128
 Db 124 LNPVCK 130
 RESULT 11
 AAR38228
 ID AAR38228 standard; protein; 130 AA.
 XX AC AAR38228;
 XX 25-MAR-2003 (updated)
 DT 01-OCT-1993 (first entry)
 XX Sequence of polypeptide chain of alboaggregin B1 (AL-B2).
 DE Venom; snake; platelet-binding protein.
 XX Trimeresurus albolabris.
 XX WO9311151-Al.
 XX 10-JUN-1993.
 XX 01-DEC-1992; 92WO-US10344.
 XX 03-DEC-1991; 91US-0803630.
 PR 05-JUN-1992; 92US-0893929.
 XX (UTEM) UNIV TEMPLE.
 PA Kirby EP, Peng M;
 PI WPI; 1993-196991/24.
 XX New platelet-binding proteins obtained from snake venom - inhibit
 PT binding of von Willebrand factor to platelet membrane
 PT glycoprotein IB, useful in therapy, diagnosis and surgery
 XX Claim 8; Page 48; 74pp; English.
 XX Alboaggregin B1 and B2 (collectively "AL-B"), may be obtained in
 CC substantially pure chemical form from venom. Each of AL-B1 and -B2
 CC comprise an about 23 kDa protein formed by an about 17 kDa
 CC polypeptide chain crosslinked to an about 14 kDa polypeptide chain
 CC by one or more interchain disulfide bonds. Each of the two
 CC polypeptide chains of AL-B1 and -B2 show strong homology to two of
 CC the chains of AL-A. AL-B1 and -B2 each comprise two non-identical
 CC polypeptide chains.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX SQ Sequence 130 AA;
 Query Match 55.2%; Score 395; DB 14; Length 130;
 Best Local Similarity 54.7%; Pred. No. 3.5e-34;
 Matches 70; Conservative 22; Mismatches 34; Indels 2; Gaps 1;
 QY 1 DCSSDWSYEGHCYKVFQKSTWTDASFCYKQVNGGHLVSISSGSEADFGVQLIAQKIK 60

Db 1 DCPDSSSEKQYCIQVVKELKTWEDAEFCSEQANDGHVLSIESYREAVFVALLSENK 60
 QY 61 SAKIHVIGLRAONKEKQCSISWSDGSSISKENWIEESKKCLGVHIETGPHKWFNFCYCE 120
 Db 61 --KYHWIGLGVONKQGCSSSEWSDGSSSVSYENLVKPNPKCFVLKKESEFTWSNVCYCE 118
 QY 121 QODPFVCE 128
 Db 119 QKHIFVCK 126

RESULT 12
 ID AAE20181
 XX AAE20181 standard; Protein; 158 AA.
 AC AAE20181;

DT 18-JUN-2002 (first entry)
 XX Pigmy rattlesnake venom gland protein, Zank5.
 DE Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;
 KW platelet aggregation; Zank5.
 XX Sistrurus miliarius.

OS Key Location/Qualifiers
 FH Peptide 1..23
 FT /label= Signal_peptide
 FT Protein 24..158
 FT /label= Mature_Zank5_protein
 FT Disulfide-bond 27..38
 FT Disulfide-bond 55..152
 FT Disulfide-bond 127..144

XX WO200214364-A2.

XX 21-FEB-2002.

XX 13-AUG-2001; 2001WO-US95310.

XX 14-AUG-2000; 2000US-225072P.

XX 14-AUG-2000; 2000US-225087P.

XX 15-AUG-2000; 2000US-225489P.

XX 15-AUG-2000; 2000US-225490P.

XX 20-DEC-2000; 2000US-356997P.

XX (ZYMO) ZYMOGENETICS INC.

XX Sheppard PO, Bishop PD;

XX WPI; 2002-269180/31.

XX N-PSDB; AAD32059.

XX New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which
 affect blood coagulation and platelet aggregation system, useful in
 therapy and diagnostics, or as tools in the study of genetics or
 molecular biology

XX Claim 3; Page 3; 79pp; English.

XX The invention relates to new pigmy rattlesnake (Sistrurus miliarius)
 venom gland proteins, which affect blood coagulation and platelet
 aggregation system. The polypeptides, which affect blood coagulation and
 platelet aggregation system, are useful in therapy and diagnostics. The
 polypeptides are also useful as an educational tool in laboratory
 practical kits for courses related to genetics and molecular biology,
 protein chemistry and antibody production and analysis. The
 polynucleotide or polypeptide can be used as standards or as unknowns
 for testing purposes. The polypeptides are also useful in identifying
 proteins by western blotting, protein purification, determining the
 weight of expressed polypeptides as a ratio to total protein expressed,

CC identifying peptide cleavage sites, coupling amino and carboxyl terminal
 CC tags, mass spectrometry, circular dichroism to determine conformation or
 CC affinity chromatography columns to purify the protein, cloning or
 CC sequencing. The present sequence is Sistrurus miliarius venom gland
 CC protein, Zank5.
 XX

SQ Sequence 158 AA;

Query Match 54.1%; Score 387; DB 23; Length 158;

Best Local Similarity 52.3%; Pred. No. 3.1e-33;

Matches 67; Conservative 25; Mismatches 36; Indels 0; Gaps 0;

QY 1 DCSSDSSVEGHCYKVFQSKTWTDAESFCTKQVNGHLSIESGEADFVGLIAQKIK 60

Db 26 NCPSCGFAYDQYCYRVIKRLKTWDDAEFCSEQANGHLSVENEAEVFLAQLVAANIK 85

QY 61 SAKIHVIGLRAONKEKQCSISWSDGSSISKENWIEESKKCLGVHIETGPHKWFNFCYCE 120

Db 86 QNYYVWIGLRIQNGQCCSTKWSGSSVSYENLVKSHSKKCFGLKKEFTLQWNTDCE 145

QY 121 QODPFVCE 128

Db 146 EKNLFCVK 153

RESULT 13

AAR72235

ID AAR72235 standard; peptide; 132 AA.

AC AAR72235;

XX 25-MAR-2003 (updated)

DT 04-DEC-1995 (first entry)

XX Vipera palestinae venom derived peptide.

XX Vipera palestinae; venom; antithrombotic agent;

KW platelet binding; von Willebrand factor.

XX Vipera palestinae.

XX WO9509183-A1.

XX 06-APR-1995.

XX 27-SEP-1994; 94WO-JP01583.

XX 28-SEP-1993; 93JP-0241666.

XX 14-JAN-1994; 94JP-0002691.

XX 10-JUN-1994; 94JP-0128518.

XX (AJIN) AJINOMOTO KK.

XX Fukuchi N, Ishii K, Kaida K, Kobayashi T;

XX WPI; 1995-147392/19.

XX Peptide derived from Vipera palestinae venom - inhibits binding
 of platelets to von Willebrand factor, useful as antithrombotic
 agent

XX Claim 5; Page 19; 37pp; Japanese.

XX AAR72235 is a Vipera palestinae venom derived peptide. A peptide
 compn. comprising AAR72235 and AAR72236 is claimed, it inhibits the
 binding of platelets to von Willebrand factors and can therefore
 be used as an anti-thrombotic agent.
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 132 AA;

Query Match 52.0%; Score 372; DB 16; Length 132;

Best Local Similarity 49.6%; Pred. No. 9.9e-32;

121 DKIPFICKS 129

Qy 61 SAKI --HVWIGLRAQKQCSIEWSDGSSISKENWIEESKCLGVHIEIGFHKWENFY 118

search completed: DE

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:50:40 ; Search time 28.0114 Seconds
(without alignments)
194.853 Million cell updates/sec

Title: US-09-938-114-2

Perfect score: 715

Sequence: 1 DCSDSWSYEGHCYKFKQS.....GPHKWFYCEQDPFVCEA 129

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:**

1: /cgn2_6/ptodata/1/iaa/5A COMB.pep:**

2: /cgn2_6/ptodata/1/iaa/5B COMB.pep:**

3: /cgn2_6/ptodata/1/iaa/6A COMB.pep:**

4: /cgn2_6/ptodata/1/iaa/6B COMB.pep:**

5: /cgn2_6/ptodata/1/iaa/PTUS COMB.pep:**

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	715	100.0	129	4	US-09-058-740-2
2	425	59.4	131	1	US-07-893-929A-1
3	425	59.4	131	5	PCT-US92-10344-1
4	406	56.8	132	1	US-07-893-929A-5
5	406	56.8	132	5	PCT-US92-10344-5
6	404	56.5	134	1	US-07-893-929A-2
7	404	56.5	134	5	PCT-US92-10344-2
8	395	55.2	130	1	US-07-893-929A-7
9	395	55.2	130	5	PCT-US92-10344-7
10	309	43.2	133	1	US-07-893-929A-9
11	309	43.2	133	5	PCT-US92-10344-9
12	292	40.8	123	1	US-07-893-929A-4
13	292	40.8	123	5	PCT-US92-10344-4
14	286.5	40.1	126	2	US-08-612-840A-2
15	286.5	40.1	126	2	US-08-612-840A-8
16	274.5	38.4	127	1	US-07-614-443A-1
17	274.5	38.4	127	1	US-08-294-859-1
18	274.5	38.4	127	1	US-08-481-676-1
19	259	36.2	117	1	US-07-614-443A-2
20	259	36.2	117	1	US-08-294-859-2
21	259	36.2	117	1	US-08-481-676-2
22	246.5	34.5	107	1	US-07-893-929A-6
23	246.5	34.5	107	5	PCT-US92-10344-6
24	230.5	32.2	123	1	US-07-893-929A-10
25	230.5	32.2	123	5	PCT-US92-10344-10
26	222	31.0	125	1	US-07-893-929A-3
27	222	31.0	125	5	PCT-US92-10344-3

28 214 29.9 50 1 US-07-614-443A-5 Sequence 5, Appli
29 214 29.9 50 1 US-08-294-859-5 Sequence 5, Appli
30 214 29.9 50 1 US-08-481-676-5 Sequence 5, Appli
31 214 29.9 128 1 US-07-893-929A-8 Sequence 8, Appli
32 214 29.9 128 5 PCT-US92-10344-8 Sequence 8, Appli
33 201.5 28.2 1479 3 US-08-840-062-4 Sequence 4, Appli
34 197.5 27.6 1479 3 US-08-840-062-2 Sequence 2, Appli
35 192.5 26.9 166 2 US-08-729-103-4 Sequence 4, Appli
36 186.5 26.1 174 2 US-08-401-530A-5 Sequence 5, Appli
37 186.5 26.1 174 2 US-08-709-662-5 Sequence 5, Appli
38 184.5 25.8 1257 1 US-08-340-428B-49 Sequence 49, Appli
39 183 25.6 174 1 US-07-778-156-7 Sequence 7, Appli
40 183 25.6 174 2 US-08-822-261-4 Sequence 4, Appli
41 183 25.6 174 2 US-08-422-166-7 Sequence 7, Appli
42 183 25.6 174 4 US-09-226-852-4 Sequence 4, Appli
43 183 25.6 175 2 US-08-464-637-2 Sequence 2, Appli
44 183 25.6 175 2 US-08-401-530A-4 Sequence 4, Appli
45 183 25.6 175 2 US-08-709-662-4 Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-058-740-2
; Sequence 2, Application US/09058740
; Patent No. 6489451
; GENERAL INFORMATION:
; APPLICANT: Chun Wang, Benjamin XV Li, Xin Cheng,
; Jing Liu, Li-Wen Niu, Wan-Zhi Huang,
; Zhen-Yu Xu, Dan Luo, Lian-Di Kang,
; Jin-Guo Ding, Fang Rong, Yan Liu and
; Hui-Ran Chen
; TITLE OF INVENTION: AN ANTITHROMBOSIS ENZYME FROM THE SNAKE
; VENOM OF AGKISTRODON ACUTUS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; SUITE: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FASTSEQ for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/058,740
; FILING DATE: 10-Apr-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Chen, Anthony C.
; REGISTRATION NUMBER: 38,673
; REFERENCE/DOCKET NUMBER: 233/298
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 129 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: Peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-058-740-2

Query Match 100.0%; Score 715; DB 4; Length 129;
Best Local Similarity 100.0%; Pred. No. 7.1e-71;
Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DCSDWSSYEGHCYKVFQKSTWTDASFCTKVQNGHLSIESSEADFGVQLIAQKIK 60
Db 1 DCSDWSSYEGHCYKVFQKSTWTDASFCTKVQNGHLSIESSEADFGVQLIAQKIK 60
Qy 61 SAKIHVWIGLRAQNKKEKCSIEWSDGSSISKENWIEESKKCLGVHIETGPHKWNFYCE 120
Db 61 SAKIHVWIGLRAQNKKEKCSIEWSDGSSISKENWIEESKKCLGVHIETGPHKWNFYCE 120
Qy 121 QDDPFVCEA 129
Db 121 QDDPFVCEA 129

RESULT 2

US-07-893-929A-1

; Sequence 1, Application US/07893929A

; Patent No. 5336667

; GENERAL INFORMATION:

; APPLICANT: Kirby, Edward P.

; APPLICANT: Peng, Man-ling

; TITLE OF INVENTION: Alboaggregins: Platelet

; TITLE OF INVENTION: Agonists Which Bind To Platelet

; TITLE OF INVENTION: Membrane Glycoprotein Ib

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Temple University - Of The Common-

; ADDRESSEE: wealth System of Higher Education

; STREET: 406 University Services Building

; CITY: Philadelphia

; STATE: Pennsylvania

; COUNTRY: U.S.A.

; ZIP: 19122

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: WordPerfect 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/893,929A

; FILING DATE: 19920605

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/803,630

; FILING DATE: December 3, 1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Monaco, Daniel A.

; REGISTRATION NUMBER: 30,480

; REFERENCE/DOCKET NUMBER: 6056-126 (CIP) 1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (215) 568-8383

; TELEFAX: (215) 568-5549

; TELEX: No. 5336667E

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 131 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: linear

US-07-893-929A-1

Query Match 59.4%; Score 425; DB 1; Length 131;
Best Local Similarity 57.0%; Pred. No. 4e-39;
Matches 73; Conservative 22; Mismatches 33; Indels 0; Gaps 0;

Qy 1 DCSDWSSYEGHCYKVFQKSTWTDASFCTKVQNGHLSIESSEADFGVQLIAQKIK 60
Db 1 DCPDSSYDQCYRVFKRLQWEDARFCSEQANDGHLVSTESAGEADFTQLVSENIR 60

Qy 61 SAKIHVWIGLRAQNKKEKCSIEWSDGSSISKENWIEESKKCLGVHIETGPHKWNFYCE 120
Db 61 SEKHYVWIGLRVQGGQCCSSEWSGDSSVHYDNLQEKTRKCYGLEKRAEFTWNSVYCG 120
Qy 121 QDDPFVCE 128
Db 121 HEYFVCK 128

RESULT 3

PCT-US92-10344-1

; Sequence 1, Application PC/TUS9210344

; GENERAL INFORMATION:

; APPLICANT: Kirby, Edward P.

; APPLICANT: Peng, Man-ling

; TITLE OF INVENTION: Alboaggregins: Platelet

; TITLE OF INVENTION: Agonists Which Bind To Platelet

; TITLE OF INVENTION: Membrane Glycoprotein Ib

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Temple University - Of The Common-

; ADDRESSEE: wealth System of Higher Education

; STREET: 406 University Services Building

; CITY: Philadelphia

; STATE: Pennsylvania

; COUNTRY: U.S.A.

; ZIP: 19122

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: WordPerfect 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US92/10344

; FILING DATE: 19921201

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/803,630

; FILING DATE: December 3, 1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Monaco, Daniel A.

; REGISTRATION NUMBER: 30,480

; REFERENCE/DOCKET NUMBER: 6056-126 (CIP) 1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (215) 568-8383

; TELEFAX: (215) 568-5549

; TELEX: None

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 131 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: linear

PCT-US92-10344-1

Query Match 59.4%; Score 425; DB 5; Length 131;
Best Local Similarity 57.0%; Pred. No. 4e-39;
Matches 73; Conservative 22; Mismatches 33; Indels 0; Gaps 0;

Qy 1 DCSDWSSYEGHCYKVFQKSTWTDASFCTKVQNGHLSIESSEADFGVQLIAQKIK 60
Db 1 DCPDSSYDQCYRVFKRLQWEDARFCSEQANDGHLVSTESAGEADFTQLVSENIR 60

Qy 61 SAKIHVWIGLRAQNKKEKCSIEWSDGSSISKENWIEESKKCLGVHIETGPHKWNFYCE 120
Db 61 SEKHYVWIGLRVQGGQCCSSEWSGDSSVHYDNLQEKTRKCYGLEKRAEFTWNSVYCG 120
Qy 121 QDDPFVCE 128
Db 121 HEYFVCK 128

RESULT 4

US-07-893-929A-5

```

; Sequence 5, Application US/078993929A
; Patent No. 5336667
; GENERAL INFORMATION:
; APPLICANT: Kirby, Edward P.
; APPLICANT: Peng, Man-ling
; TITLE OF INVENTION: Alboaggregins: Platelet
; TITLE OF INVENTION: Agonists Which Bind To Platelet
; TITLE OF INVENTION: Membrane Glycoprotein Ib
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Temple University - Of The Common-
; ADDRESSEE: wealth System of Higher Education
; STREET: 406 University Services Building
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/893,929A
; FILING DATE: 19920605
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/803,630
; FILING DATE: December 3, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 6056-126 (CIP) 1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: No. 5336667e
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 132 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; US-07-893-929A-5

Query Match 56.8%; Score 406; DB 1; Length 132;
Best Local Similarity 54.7%; Pred. No. 4.8e-37;
Matches 70; Conservative 23; Mismatches 35; Indels 0; Gaps 0;

QY 1 DCSDDSSSYEGHCYKVFQSKTWTDAESFCTKQVNGHLSIESSGRADFVGQLIAQKIK 60
Db 1 DCPDSSSSFYQCYQVIVKELTADAKFCSEQANDGHLVSIESYREAVFVAELLSNVK 60
QY 61 SAKHWIWLGRANQKQCSLEWSDGSSISKENWIEESKKCLGVHLETGPHKWFVCE 120
Db 61 TTKYNWIGLSVQNKEQCSSEWSDGSSVXYENLIKPNPKCFVKKKESEPRFTWSNVCE 120
QY 121 QQDPFVCE 128
Db 121 QKHIFMCK 128

RESULT 5
PCT-US92-10344-5
; Sequence 5, Application PC/TUS9210344
; GENERAL INFORMATION:
; APPLICANT: Kirby, Edward P.
; APPLICANT: Peng, Man-ling
; TITLE OF INVENTION: Alboaggregins: Platelet
; TITLE OF INVENTION: Agonists Which Bind To Platelet
; TITLE OF INVENTION: Membrane Glycoprotein Ib
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Temple University - Of The Common-

```

```
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/893,929A
; FILING DATE: 19920605
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/803,630
; FILING DATE: December 3, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 6056-126 (CIP) 1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: No. 5336667e
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 134 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
;
; Query Match 56.5%; Score 404; DB 1; Length 134;
; Best Local Similarity 55.9%; Pred. No. 8.1e-37;
; Matches 71; Conservative 22; Mismatches 34; Indels 0; Gaps 0;
;
QY 2 CSSDSSYEGHCYKVFQSKTWTDAESFCTKQVNGHLSVIESGSEADFGQLIAQKIKS 61
DB 4 CLPGWSAYDQCYRVFNEPKNWEDAEFCQAQDSGHLVSIETMGADFVAQLISENIQS 63
QY 62 AKIHWIGLRAONKEKQCSIEWSGSSISKENWIEESKCLGVHIETGPHKWFNFCYCEQ 121
DB 64 KEHYVWIGLVQNKREQQCSSEWSDGSSVYENLIKLYMRKCGALEQESGFRKWINLGCIO 123
QY 122 QDPFVCE 128
DB 124 LNPFCVK 130
;
; US-07-893-929A-2
;
; Query Match 56.5%; Score 404; DB 1; Length 134;
; Best Local Similarity 55.9%; Pred. No. 8.1e-37;
; Matches 71; Conservative 22; Mismatches 34; Indels 0; Gaps 0;
;
QY 2 CSSDSSYEGHCYKVFQSKTWTDAESFCTKQVNGHLSVIESGSEADFGQLIAQKIKS 61
DB 4 CLPGWSAYDQCYRVFNEPKNWEDAEFCQAQDSGHLVSIETMGADFVAQLISENIQS 63
QY 62 AKIHWIGLRAONKEKQCSIEWSGSSISKENWIEESKCLGVHIETGPHKWFNFCYCEQ 121
DB 64 KEHYVWIGLVQNKREQQCSSEWSDGSSVYENLIKLYMRKCGALEQESGFRKWINLGCIO 123
QY 122 QDPFVCE 128
DB 124 LNPFCVK 130
;
; RESULT 7
; PCT-US92-10344-2
; Sequence 2, Application PC/TUS9210344
; GENERAL INFORMATION:
; APPLICANT: Kirby, Edward P.
; TITLE OF INVENTION: Alboaggregins: Platelet
; TITLE OF INVENTION: Agonists Which Bind To Platelet
; TITLE OF INVENTION: Membrane Glycoprotein Ib
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Temple University - Of The Common-
; ADDRESS: wealth System of Higher Education
; STREET: 406 University Services Building
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/10344
; FILING DATE: 19921201
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/803,630
; FILING DATE: December 3, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 6056-126 (CIP) 1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: No. 5336667e
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 130 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
```

```
; REFERENCE/DOCKET NUMBER: 6056-126 (CIP) 1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: None
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 134 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
;
; PCT-US92-10344-2
;
; Query Match 56.5%; Score 404; DB 5; Length 134;
; Best Local Similarity 55.9%; Pred. No. 8.1e-37;
; Matches 71; Conservative 22; Mismatches 34; Indels 0; Gaps 0;
;
QY 2 CSSDSSYEGHCYKVFQSKTWTDAESFCTKQVNGHLSVIESGSEADFGQLIAQKIKS 61
DB 4 CLPGWSAYDQCYRVFNEPKNWEDAEFCQAQDSGHLVSIETMGADFVAQLISENIQS 63
QY 62 AKIHWIGLRAONKEKQCSIEWSGSSISKENWIEESKCLGVHIETGPHKWFNFCYCEQ 121
DB 64 KEHYVWIGLVQNKREQQCSSEWSDGSSVYENLIKLYMRKCGALEQESGFRKWINLGCIO 123
QY 122 QDPFVCE 128
DB 124 LNPFCVK 130
;
; RESULT 8
; US-07-893-929A-7
; Sequence 7, Application US/07893929A
; Patent No. 5336667
; GENERAL INFORMATION:
; APPLICANT: Kirby, Edward P.
; TITLE OF INVENTION: Alboaggregins: Platelet
; TITLE OF INVENTION: Agonists Which Bind To Platelet
; TITLE OF INVENTION: Membrane Glycoprotein Ib
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Temple University - Of The Common-
; ADDRESS: wealth System of Higher Education
; STREET: 406 University Services Building
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/893,929A
; FILING DATE: 19920605
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/803,630
; FILING DATE: December 3, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 6056-126 (CIP) 1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: No. 5336667e
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 130 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
```


US-07-893-929A-7

Query Match 55.2%; Score 395; DB 1; Length 130;
 Best Local Similarity 54.7%; Pred. No. 7.6e-36;
 Matches 70; Conservative 22; Mismatches 34; Indels 2; Gaps 1;

QY 1 DCSSDWSYEGHCYKVKQSKTWTDAESFCTKQVNGGHLVSISSGSEADFGVQLIAQKIK 60
 DB 1 DCPDMSWFKQYQIVKELKTWDAERFCSEQANDGHLVSISSYREAVFVALLSENK 60

QY 61 SAKIHVWIGLRAQNKQCSIEWSDGSSISKENWIEESKCLGVHETGFHKWENFYCE 120
 DB 61 --KYHVWIGLSVQNKQCSSEWSDGSSVSYENLVKPNPKCFVLKKESEFKTWSNVCE 118

QY 121 QODPFVCE 128
 DB 119 QKHFMCK 126

RESULT 9

PCT-US92-10344-7

; Sequence 7, Application PC/TUS9210344
 ; GENERAL INFORMATION:
 ; APPLICANT: Kirby, Edward P.
 ; TITLE OF INVENTION: Alboagregins: Platelet
 ; TITLE OF INVENTION: Agonists Which Bind To Platelet
 ; TITLE OF INVENTION: Membrane Glycoprotein Ib
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Temple University - Of The Common-
 ; ADDRESSEE: wealth System of Higher Education
 ; STREET: 406 University Services Building
 ; CITY: Philadelphia
 ; STATE: Pennsylvania
 ; COUNTRY: U.S.A.
 ; ZIP: 19122

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
 ; COMPUTER: IBM PS/2
 ; OPERATING SYSTEM: MS-DOS
 ; SOFTWARE: WordPerfect 5.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US92/10344
 ; FILING DATE: 19921201
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/803,630
 ; FILING DATE: December 3, 1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Monaco, Daniel A.
 ; REGISTRATION NUMBER: 30,480
 ; REFERENCE/DOCKET NUMBER: 6056-126 (CIP) 1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (215) 568-8383
 ; TELEFAX: (215) 568-5549
 ; TELEX: None
 ; INFORMATION FOR SEQ ID NO: 7:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 130 amino acids
 ; TYPE: AMINO ACID
 ; TOPOLOGY: linear

PCT-US92-10344-7

Query Match 55.2%; Score 395; DB 5; Length 130;
 Best Local Similarity 54.7%; Pred. No. 7.6e-36;
 Matches 70; Conservative 22; Mismatches 34; Indels 2; Gaps 1;

QY 1 DCSSDWSYEGHCYKVKQSKTWTDAESFCTKQVNGGHLVSISSGSEADFGVQLIAQKIK 60
 DB 1 DCPDMSWFKQYQIVKELKTWDAERFCSEQANDGHLVSISSYREAVFVALLSENK 60

QY 61 SAKIHVWIGLRAQNKQCSIEWSDGSSISKENWIEESKCLGVHETGFHKWENFYCE 120

DB 61 --KYHVWIGLSVQNKQCSSEWSDGSSVSYENLVKPNPKCFVLKKESEFKTWSNVCE 118
 QY 121 QODPFVCE 128
 DB 119 QKHFMCK 126

RESULT 10
 US-07-893-929A-9
 ; Sequence 9, Application US/07893929A
 ; Patent No. 5336667
 ; GENERAL INFORMATION:
 ; APPLICANT: Kirby, Edward P.
 ; TITLE OF INVENTION: Alboagregins: Platelet
 ; TITLE OF INVENTION: Agonists Which Bind To Platelet
 ; TITLE OF INVENTION: Membrane Glycoprotein Ib
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Temple University - Of The Common-
 ; ADDRESSEE: wealth System of Higher Education
 ; STREET: 406 University Services Building
 ; CITY: Philadelphia
 ; STATE: Pennsylvania
 ; COUNTRY: U.S.A.
 ; ZIP: 19122

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
 ; COMPUTER: IBM PS/2
 ; OPERATING SYSTEM: MS-DOS
 ; SOFTWARE: WordPerfect 5.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/893,929A
 ; FILING DATE: 19920605
 ; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/803,630
 ; FILING DATE: December 3, 1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Monaco, Daniel A.
 ; REGISTRATION NUMBER: 30,480
 ; REFERENCE/DOCKET NUMBER: 6056-126 (CIP) 1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (215) 568-8383
 ; TELEFAX: (215) 568-5549
 ; TELEX: No. 5336667e
 ; INFORMATION FOR SEQ ID NO: 9:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 133 amino acids
 ; TYPE: AMINO ACID
 ; TOPOLOGY: linear

US-07-893-929A-9

Query Match 43.2%; Score 309; DB 1; Length 133;
 Best Local Similarity 47.3%; Pred. No. 2e-26;
 Matches 62; Conservative 24; Mismatches 39; Indels 6; Gaps 4;

QY 1 DCSSDWSYEGHCYKVKQSKTWTDAESFCTKQVNGGHLVSISSGSEADFGVQLIAQKIK 60
 DB 3 DCLSGWSFYEGHCYQLFR-LKTWDEAKYC-NQWDGHLVSISSNAKAEFVAQLISRKLP 60

QY 61 SAKI--HVWIGLRAQNKQCSIEWSDGSSISKENWIEESKCLGVHETGFHKWENFY 118
 DB 61 KSAIEDRVWIGLDRSKREQCGLHWTDSNPFVYHVV--PPTKCFVLERTEFKWIADV 118

QY 119 CQODPFVCEA 129
 DB 119 CEFKFFPVCKA 129

RESULT 11
 PCT-US92-10344-9

```
; Sequence 9, Application PC/TUS9210344
; GENERAL INFORMATION:
; APPLICANT: Kirby, Edward P.
; TITLE OF INVENTION: Alboagregins: Platelet
; TITLE OF INVENTION: Agonists Which Bind To Platelet
; TITLE OF INVENTION: Membrane Glycoprotein Ib
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Temple University - Of The Common-
; ADDRESSEE: wealth System of Higher Education
; STREET: 406 University Services Building
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/10344
; FILING DATE: 19921201
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/803,630
; FILING DATE: December 3, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 6056-126 (CIP) 1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: None
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 133 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
;
PCT-US92-10344-9
Query Match 43.2%; Score 309; DB 5; Length 133;
Best Local Similarity 47.3%; Pred. No. 2e-26; Mismatches 39; Indels 6; Gaps 4;
Matches 62; Conservative 24;

Qy 1 DCSDWSYEGHCYKVFQKSTWTDAESFCTKQVNGHLSVTSSEGEADFGQLIAQKIK 60
Db 3 DCLGWSFYEGHCYQLFR-LKTWDEAKYC-NQWDGGHLSVTSSENAKAEFVAQLISRLKP 60

Qy 61 SAKI--HWIGIRAQNKQKCSIEWSGSSISKENWIEESKCLGVHIETGFHKWENFY 118
Db 61 KSAIEDRWIGLRDSKREQCGLWTDNSFVHYEHV--PPTKCFVLRQTEFRKWI 118

Qy 119 CEQQDPFVCEA 129
Db 119 CEFKPFVCKA 129

RESULT 12
US-07-893-929A-4
; Sequence 4, Application US/07893929A
; Patent No. 5336667
; GENERAL INFORMATION:
; APPLICANT: Kirby, Edward P.
; APPLICANT: Peng, Man-ling
; TITLE OF INVENTION: Alboagregins: Platelet
; TITLE OF INVENTION: Agonists Which Bind To platelet
; TITLE OF INVENTION: Membrane Glycoprotein Ib
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Temple University - Of The Common-
; ADDRESSEE: wealth System of Higher Education
; STREET: 406 University Services Building
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/10344
; FILING DATE: 19921201
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/803,630
; FILING DATE: December 3, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 6056-126 (CIP) 1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: None
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 133 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
;
US-07-893-929A-4
Query Match 40.8%; Score 292; DB 1; Length 123;
Best Local Similarity 46.9%; Pred. No. 1.3e-24;
Matches 60; Conservative 18; Mismatches 42; Indels 8; Gaps 5;

Qy 1 DCSDWSYEGHCYKVFQKSTWTDAESFCTKQVNGHLSVTSSEGEADFGQLIAQKIK 60
Db 1 DCPDWSYEGHCYKVFQKSTWTDAESFCTKQVNGHLSVTSSEGEADFGQLIAQKIK 59

Qy 61 SAKIHWIGIRAQNKQKCSIEWSGSSISKENWIEESKCLGVHIETGFHKWENFY 120
Db 60 NANL-VWIGL--SNLWNCNSQWSDGTXYLDYKXW--REQFECL--VSRTTNNEWLSMDCS 112

Qy 121 QQDPFVCE 128
Db 113 STHSFVCE 120

RESULT 13
PCT-US92-10344-4
; Sequence 4, Application PC/TUS9210344
; GENERAL INFORMATION:
; APPLICANT: Kirby, Edward P.
; APPLICANT: Peng, Man-ling
; TITLE OF INVENTION: Alboagregins: Platelet
; TITLE OF INVENTION: Agonists Which Bind To Platelet
; TITLE OF INVENTION: Membrane Glycoprotein Ib
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Temple University - Of The Common-
; ADDRESSEE: wealth System of Higher Education
; STREET: 406 University Services Building
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
```

FILING DATE: 22-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: No. 5856126man F. Oblon
REGISTRATION NUMBER: 24,618
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 126 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-612-840A-2

Query Match 40.1%; Score 286.5; DB 2; Length 126;
Best Local Similarity 43.0%; Pred. No. 5.5e-24;
Matches 55; Conservative 22; Mismatches 42; Indels 9; Gaps 2;

QY 1 DCSDWSSYEGHCYKVFQSKTWTDAESFCTKVQNGHLYSISSGADPVGQLIAQKIK 60
Db 3 ECPGSWSYDRYCYRPFKQEMTWADAERFCSEQAAGHLLSVETALEASFDVNLVANKE 62
QY 61 SAKIHWITGLRAQNKCSTIWDGSGSISKENWIEESKKCLGVHIETGFHKWFNPFCE 120
Db 63 YLTRYIWIGLRVQNKGPC-----SSISYENLV--DPFECPMVSRTDLREWFVKVDCE 113
QY 121 QODPFVCE 128
Db 114 QQSHPICK 121

RESULT 15
US-08-612-840A-8
Sequence 8, Application US/08612840A
Patent No. 5856126
GENERAL INFORMATION:
APPLICANT: FUKUCHI, Naoyuki
APPLICANT: YAMAMOTO, Hiroshi
APPLICANT: NAGANO, Mitsuyo
APPLICANT: KITO, Morikazu
APPLICANT: TANAKA, Akiko
APPLICANT: ISHII, Koichi
APPLICANT: KOBAYASHI, Tsuyoshi
APPLICANT: YOSHIMOTO, Ryota
TITLE OF INVENTION: PEPTIDE HAVING ANTI-THROMBUS ACTIVITY AND
TITLE OF INVENTION: METHOD OF PRODUCING THE SAME
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Oblon, Spivak, McClelland, Maier & Neustadt, P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,840A
FILING DATE: 20-MAR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-236975
FILING DATE: 22-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: No. 5856126man F. Oblon
REGISTRATION NUMBER: 24,618
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220

FILING DATE: 19921201
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/803,630
FILING DATE: December 3, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A.
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 6056-126 (CIP) 1
TELEPHONE: (215) 568-8383
TELEFAX: (215) 568-5549
TELEX: None
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 123 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
PCT-US92-10344-4

Query Match 40.8%; Score 292; DB 5; Length 123;
Best Local Similarity 46.9%; Pred. No. 1.3e-24;
Matches 60; Conservative 18; Mismatches 42; Indels 8; Gaps 5;

QY 1 DCSDWSSYEGHCYKVFQSKTWTDAESFCTKVQNGHLYSISSGADPVGQLIAQKIK 60
Db 1 DCPDSSSYEGHCYRVFNPNWADAERFKTCQQKHSHLVFSQSSSEADFPVQM-TRPIL 59
QY 61 SAKIHWITGLRAQNKCSTIWDGSGSISKENWIEESKKCLGVHIETGFHKWFNPFCE 120
Db 60 NANL-VWIGL--SNLWNCNSQWSDGTXYLDYKXW--REQECL--VSRRTTNNEWLMDCS 112
QY 121 QODPFVCE 128
Db 113 STHSFVCE 120

RESULT 14
US-08-612-840A-2
Sequence 2, Application US/08612840A
Patent No. 5856126
GENERAL INFORMATION:
APPLICANT: FUKUCHI, Naoyuki
APPLICANT: YAMAMOTO, Hiroshi
APPLICANT: NAGANO, Mitsuyo
APPLICANT: KITO, Morikazu
APPLICANT: TANAKA, Akiko
APPLICANT: ISHII, Koichi
APPLICANT: KOBAYASHI, Tsuyoshi
APPLICANT: YOSHIMOTO, Ryota
TITLE OF INVENTION: PEPTIDE HAVING ANTI-THROMBUS ACTIVITY AND
TITLE OF INVENTION: METHOD OF PRODUCING THE SAME
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Oblon, Spivak, McClelland, Maier & Neustadt, P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,840A
FILING DATE: 20-MAR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-236975

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:54:55 ; Search time 53.0743 Seconds
(without alignments)
452.044 Million cell updates/sec

Title: US-09-938-114-2

Perfect score: 715

Sequence: 1 DCSWSSVEGHCYKVFQKS.....GPHKWFYCEQDPFVCEA 129

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1:	/cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2:	/cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6:	/cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8:	/cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10:	/cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12:	/cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13:	/cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14:	/cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15:	/cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16:	/cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17:	/cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18:	/cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	715	100.0	129	11	US-09-938-114-2
2	477	66.7	152	10	US-09-929-230-5
3	477	66.7	152	12	US-10-226-420-5
4	434	60.7	144	10	US-09-929-230-8
5	434	60.7	144	12	US-10-226-420-8
6	387	54.1	158	10	US-09-929-230-11
7	387	54.1	158	12	US-10-226-420-11
8	286.5	40.1	149	10	US-09-969-763-3
9	271.5	38.0	126	10	US-09-969-763-1
10	264	36.9	151	10	US-09-929-230-2
11	264	36.9	151	12	US-10-226-420-2
12	223.5	31.3	110	10	US-09-969-763-10
13	201.5	28.2	154	9	US-09-764-870-439
14	201.5	28.2	154	15	US-10-125-540-439
15	201.5	28.2	166	12	US-09-997-003-30

16	201.5	28.2	166	12	US-09-997-003-43
17	201.5	28.2	174	9	US-09-925-301-1182
18	201.5	28.2	250	9	US-09-764-870-280
19	201.5	28.2	250	15	US-10-125-540-280
20	201.5	28.2	1479	15	US-10-015-219-1738
21	201.5	28.2	1479	15	US-10-015-219-1739
22	192.5	26.9	174	9	US-09-925-297-753
23	184.5	25.6	1321	12	US-10-241-220-82
24	183	25.6	174	15	US-10-316-761-4
25	183	25.6	175	15	US-10-316-761-3
26	183	25.6	183	9	US-09-925-301-1013
27	181.5	25.4	158	9	US-09-922-217-1070
28	181.5	25.4	158	9	US-09-922-217-1077
29	181.5	25.4	158	9	US-09-922-217-1078
30	181.5	25.4	158	9	US-09-922-217-1079
31	181.5	25.4	158	9	US-09-922-217-1080
32	181.5	25.4	158	10	US-09-833-263-1070
33	181.5	25.4	158	10	US-09-833-263-1077
34	181.5	25.4	158	10	US-09-833-263-1078
35	181.5	25.4	158	10	US-09-833-263-1079
36	181.5	25.4	158	10	US-09-833-263-1080
37	181.5	25.4	158	12	US-09-525-041-2
38	181.5	25.4	158	14	US-10-025-380-1070
39	181.5	25.4	158	14	US-10-025-380-1077
40	181.5	25.4	158	14	US-10-025-380-1078
41	181.5	25.4	158	14	US-10-025-380-1079
42	181.5	25.4	158	14	US-10-025-380-1080
43	181.5	25.4	158	15	US-10-100-608B-2
44	181.5	25.4	158	15	US-10-157-031-114
45	181.5	25.4	158	15	US-10-205-823-347

ALIGNMENTS

RESULT 1

US-09-938-114-2
; Sequence 2, Application US/09938114
; Publication No. US20030022350A1
; GENERAL INFORMATION:
; APPLICANT: Chun Wang, Benjamin XY Li, Xin Cheng,
; Jing Liu, Li-Wen Niu, Wan-Zhi Huang,
; Zhen-Yu Xu, Dan Luo, Lian-Di Kang,
; Jin-Guo Ding, Fang Rong, Yan Liu and
; Hui-Ran Chen
; TITLE OF INVENTION: AN ANTITHROMBOSIS ENZYME FROM THE SNAKE
; VENOM OF AGKISTRODON ACUTUS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FASTSEQ for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/938,114
; FILING DATE: 23-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/058,740
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Chen, Anthony C.
; REGISTRATION NUMBER: 38,673
; REFERENCE/DOCKET NUMBER: 233/298

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 129 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-938-114-2

Query Match 100.0%; Score 715; DB 11; Length 129;
Best Local Similarity 100.0%; Pred. No. 9.3e-70;
Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DCSSDSSSYEGHCYKVFQKSTWTDAESFCTKQVNGHGLVSTESSGEADFGQLIAQKIK 60
DB 1 DCSSDSSSYEGHCYKVFQKSTWTDAESFCTKQVNGHGLVSTESSGEADFGQLIAQKIK 60
QY 61 SAKIHVWIGLRAQNKKEKCSIEWSDGSSISKENWIEESKKCLGVHIETGPHKWENFYCE 120
DB 61 SAKIHVWIGLRAQNKKEKCSIEWSDGSSISKENWIEESKKCLGVHIETGPHKWENFYCE 120
QY 121 QODPFVCEA 129
DB 121 QODPFVCEA 129

```

RESULT 2

```

US-09-929-230-5
; Sequence 5, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/09/929,230
; CURRENT FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 152
; TYPE: PRT
; ORGANISM: Sistrurus miliarius
US-09-929-230-5

```

```

Query Match 66.7%; Score 477; DB 10; Length 152;
Best Local Similarity 64.3%; Pred. No. 7.1e-44;
Matches 83; Conservative 16; Mismatches 30; Indels 0; Gaps 0;

QY 1 DCSSDSSSYEGHCYKVFQKSTWTDAESFCTKQVNGHGLVSTESSGEADFGQLIAQKIK 60
DB 24 DCPGSSSYDQHCYRVFKQLKTWDDAERFCSEQAEGHGLVSTESSGEAAFAVLVPENRR 83
QY 61 SAKIHVWIGLRAQNKKEKCSIEWSDGSSISKENWIEESKKCLGVHIETGPHKWENFYCE 120
DB 84 RALLYIWIGLVRQGEKQCSAKWSGSSVYENWIEESKTCGLGQQTGNVHKWNIYCG 143
QY 121 QODPFVCEA 129
DB 144 EINPFVCEA 152

```

RESULT 3

```

US-10-226-420-5
; Sequence 5, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.

```

```

; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/10/226,420
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 152
; TYPE: PRT
; ORGANISM: Sistrurus miliarius
US-10-226-420-5

```

```

Query Match 66.7%; Score 477; DB 12; Length 152;
Best Local Similarity 64.3%; Pred. No. 7.1e-44;
Matches 83; Conservative 16; Mismatches 30; Indels 0; Gaps 0;

QY 1 DCSSDSSSYEGHCYKVFQKSTWTDAESFCTKQVNGHGLVSTESSGEADFGQLIAQKIK 60
DB 24 DCPGSSSYDQHCYRVFKQLKTWDDAERFCSEQAEGHGLVSTESSGEAAFAVLVPENRR 83
QY 61 SAKIHVWIGLRAQNKKEKCSIEWSDGSSISKENWIEESKKCLGVHIETGPHKWENFYCE 120
DB 84 RALLYIWIGLVRQGEKQCSAKWSGSSVYENWIEESKTCGLGQQTGNVHKWNIYCG 143
QY 121 QODPFVCEA 129
DB 144 EINPFVCEA 152

```

RESULT 4

```

US-09-929-230-8
; Sequence 8, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/09/929,230
; CURRENT FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Sistrurus miliarius
US-09-929-230-8

```

```

Query Match 60.7%; Score 434; DB 10; Length 144;
Best Local Similarity 58.6%; Pred. No. 3e-39;
Matches 75; Conservative 20; Mismatches 33; Indels 0; Gaps 0;

QY 1 DCSSDSSSYEGHCYKVFQKSTWTDAESFCTKQVNGHGLVSTESSGEADFGQLIAQKIK 60
DB 12 DCPDWSYAYDQCYRVIKQLRTWDDAERFCSEQAEGHGLVSTESSGEAAFAVLVAENIK 71
QY 61 SAKIHVWIGLRAQNKKEKCSIEWSDGSSISKENWIEESKKCLGVHIETGPHKWENFYCE 120
DB 72 QNKYDVIWIGLRIOGEBKQCSAKWSGSSVYENWIEESKTCGLGQQTGNVHKWNIYCG 131
QY 121 QODPFVCEA 128
DB 132 QQNLFMCX 139

```

RESULT 5

```

US-10-226-420-8
; Sequence 8, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.

```

RESULT 7
US-10-226-420-11
; Sequence 11, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS

Query Match	40.1%;	Score	286.5;	DB	10;	Length	149;
Best Local Similarity	43.0%;	Pred. No.	3.1e-23;				
Matches	55;	Conservative	22;	Mismatches	42;	Indels	9;
						Gaps	2;
QY	1	DCSSDWSSYEGHCYKVFQKQKTWTDAESCTKQVNGGHLVSISSGEADYVQGIIAQKIK	60				
Db	26	ECPSGWSSYDRICYKPFQKQEMTWADAEFCSEQAGKGHLLSVETALEASFVDNVLKYANKE	85				
QY	61	SAKIHVWIGLRAQNEKQCSIEWSDCSSISKENWIEESKKCLGVHIEGTGHKWFNPFCE	120				
Db	86	YLTRYIWIQLRVNQKQPC-----SSISYENLV--DPFECFVMSRDTLRLEWFKVDCE	136				
QY	121	QQDPFPVCE	128				
Db	137	OOHSFICK	144				

```
RESULT 9
US-09-969-763-1
; Sequence 1, Application US/09969763
; Publication No. US20020198363A1
; GENERAL INFORMATION:
; APPLICANT: FUKUCHI, NAOYUKI
; APPLICANT: KITO, MORIKAZU
; APPLICANT: KAYAHARA, TAKASHI
; APPLICANT: FUTAKI, FUMIE
; APPLICANT: ISHIKAWA, KOHJI
; APPLICANT: SUZUKI, EIICHIRO
; APPLICANT: GONDOH, KEIKO
; APPLICANT: SHIMBA, NOBUHISA
; APPLICANT: YAMADA, NAOYUKI
; TITLE OF INVENTION: PROTEIN HAVING ANTITHROMBOTIC ACTIVITY AND METHOD FOR PRODUCING THE SAME
; FILE REFERENCE: 214760US0
; CURRENT APPLICATION NUMBER: US/09/969,763
; CURRENT FILING DATE: 2000-10-25
; PRIOR APPLICATION NUMBER: JP 2000-305279
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 126
; TYPE: PRT
; ORGANISM: Crotallus horridus
US-09-969-763-1

Query Match 38.0%; Score 271.5; DB 10; Length 126;
Best Local Similarity 41.4%; Pred. No. 1e-21;
Matches 53; Conservative 22; Mismatches 44; Indels 9; Gaps 2;

QY 1 DCSDWSYEGHCYKVFQKQSKTWTDAESFCTQVNGHLSVTSSEGEADFGQLIAQKIK 60
DB 3 ECPGWSSTDRYCYKVPFKQEMTWSAERFCSEQAQGHLLSVETALEASFVDNVLANK 62
QY 61 SAKIHVIGLRAONKEKQCSIEWSDGSSISKENWIEESKCLGVHIETGFHKWENFYCE 120
DB 63 YLTRYIWLGLRVQNGQPC-----SSISYENLV--DPFECFVMSVTRLRLEWFKVDCE 113
QY 121 QQDPFVCE 128
DB 114 QQHSFICK 121

RESULT 10
US-09-929-230-2
; Sequence 2, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul D.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/09/929,230
; CURRENT FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Sistrurus miliarius
US-09-929-230-2

Query Match 36.9%; Score 264; DB 10; Length 151;
Best Local Similarity 43.0%; Pred. No. 8.5e-21;
Matches 55; Conservative 18; Mismatches 49; Indels 6; Gaps 3;

QY 1 DCSDWSYEGHCYKVFQKQSKTWTDAESFCTQVNGHLSVTSSEGEADFGQLIAQKIK 60
DB 20 DCPDWSYDQHCYKVFSELKTWDAAESFCTQHRDSRLASHSEEEAFVGLKASQTIK 79
QY 61 SAKIHVIGLRAONKEKQCSIEWSDGSSISKENWIEESKCLGVHIETGFHKWENFYCE 120
DB 63 YLTRYIWLGLRVQNGQPC-----SSISYENLV--DPFECFVMSVTRLRLEWFKVDCE 113
QY 121 QQDPFVCE 128
DB 114 QQHSFICK 121
```

```
RESULT 11
US-10-226-420-2
; Sequence 2, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul D.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/10/226,420
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Sistrurus miliarius
US-10-226-420-2

Query Match 36.9%; Score 264; DB 12; Length 151;
Best Local Similarity 43.0%; Pred. No. 8.5e-21;
Matches 55; Conservative 18; Mismatches 49; Indels 6; Gaps 3;

QY 1 DCSDWSYEGHCYKVFQKQSKTWTDAESFCTQVNGHLSVTSSEGEADFGQLIAQKIK 60
DB 20 DCPDWSYDQHCYKVFSELKTWDAAESFCTQHRDSRLASHSEEEAFVGLKASQTIK 79
QY 61 SAKIHVIGLRAONKEKQCSIEWSDGSSISKENWIEESKCLGVHIETGFHKWENFYCE 120
DB 80 FTS--MWIGLK--DLWKECKWQSDDTKLDYKAWTRR--PYCTVMVVKTDRIWFENRGCE 133
QY 121 QQDPFVCE 128
DB 134 KTVSFVCK 141

RESULT 12
US-09-969-763-10
; Sequence 10, Application US/09969763
; Publication No. US20020198363A1
; GENERAL INFORMATION:
; APPLICANT: FUKUCHI, NAOYUKI
; APPLICANT: KITO, MORIKAZU
; APPLICANT: KAYAHARA, TAKASHI
; APPLICANT: FUTAKI, FUMIE
; APPLICANT: ISHIKAWA, KOHJI
; APPLICANT: SUZUKI, EIICHIRO
; APPLICANT: GONDOH, KEIKO
; APPLICANT: SHIMBA, NOBUHISA
; APPLICANT: YAMADA, NAOYUKI
; TITLE OF INVENTION: PROTEIN HAVING ANTITHROMBOTIC ACTIVITY AND METHOD FOR PRODUCING THE SAME
; FILE REFERENCE: 214760US0
; CURRENT APPLICATION NUMBER: US/09/969,763
; CURRENT FILING DATE: 2000-10-25
; PRIOR APPLICATION NUMBER: JP 2000-305279
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 10
; LENGTH: 110
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; OTHER INFORMATION: SYNTHETIC PEPTIDE
US-09-969-763-10
```


			Query Match	31.3%;	Score 223.5;	DB 10;	Length 110;
			Best Local Similarity	35.2%;	Pred. No. 1.4e-16;		
			Matches	45;	Conservative	18; Mismatches	40; Indels
						25; Gaps	1
QY	1	DCSSDWSYSRGHCVKVKSQTWTDAISGCTKQNVNGHLVSISSGEADFVGQLIAQRK	60	:	:	:	:
Dd	3	ECPSGWSSRYRCYKPKQEINTADAEKFCSEQAAGHLLSVETALEASFVDNVLVANKE	62	:	:	:	:
QY	61	SAXIHWIGLRAQNKEQCSTEWSDGSSI KENNIIEBSKKCLGVHTITGFHKKNFVCE	120	:	:	:	:
Dd	63	YLTRIWIIGLRFFPFE-----CFMVSRDTRLREWFKVDFCE	97	:	:	:	:

```

RESULT 13
US-09-764-870-439
; Sequence 439, Application US/09764870
; Patent No. US20020042386A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ14
; CURRENT APPLICATION NUMBER: US/09/764,870
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 439
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (141)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-870-439

```

```

Query Match      28.2%; Score 201.5; DB 9; Length 154;
Best Local Similarity 28.9%; Pred. No. 5.2e-14;
Matches 39; Conservative 32; Mismatches 47; Indels 17; Gaps 5;

QY      1  DCSDDSSSYEGHCYKVPKQSKTWTDAESPTCKQVNGCHLSYISSGADRVGQLIAQKIK 60
Db      11  ECEPSWQPGHCYRLQAEKRSQESKKACLR--GGGDLVSIHSMAELEFITKQKEVE 68
QY      61  SAKIHWIGLRANQKQKQCSIEWSGSSISKENW-----IEESKKCLGVHLETGFHK 113
Db      69  ----ELWIGL--NDLKLQMFNSDGLSVFTWHPPEPNFRDLSLEDCTVIINGPEG--R 120
QY      114  WENFYCEQQDPFVCE 128
Db      121  WNDSPCNOSLPISICK 135

```

```

RESULT 14
US-10-125-540-439
; Sequence 439, Application US/10125540
; Publication No. US20030059875A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ14C1
; CURRENT APPLICATION NUMBER: US/10/125,540
; CURRENT FILING DATE: 2002-04-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 439
; LENGTH: 154

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (141)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-125-540-439

Query March 28.2%; Score 201.5; DB 15; Length 154;
Best Local Similarity 28.9%; Pred. No. 5.2e-14;
Matches 39; Conservative 32; Mismatches 47; Indels 17; Gaps 5

Qy 1 DCSDDWSYEGHCYKVFQKSKTWTDAESFCTQVNGGHLVSIESSGEADFVGQLIAQKIK 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 11 ECEPSWQPFQGHCVRLQAEKRSQWSEKKACL--GGGDLVSIHSAELEFITKIQKEVE 68
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 61 SAKIHVWIGLRAQNKKEQCSIEWDGGSSISKENW-----IEESKKCLGVHIITGFHK 113
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 69 ----ELMIGL--NDLKLQMFWDSDGLSVSFTHWHFFPNFRDLSLEDCVTIWGPEG--R 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 114 WENFYCEQQDPPFYCE 128
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 WNDSPCNQSLPSICK 135
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 15
US-09-997-003-30
; Sequence 30, Application US/09997003
; Publication No. US20030203361A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA003P1
; CURRENT APPLICATION NUMBER: US/09/997,003
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: unassigned
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: PCT/US00/22157
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/148,680
; PRIOR FILING DATE: 1999-08-13
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-997-003-30

```

Query Match	28.2%;	Score	201.5;	DB	12;	Length	166;
Best Local Similarity	33.3%;	Pred. No.	5.7e-14;				
Matches	44;	Conservative	26;	Mismatches	53;	Indels	9;
						Gaps	5
QY	2	CSSDWSYEGHCYKVPKQSKTWTDAESFCYQVNGHLSVIESGHEAD	FV	COLIAQTKIS	61		
Db	36	CPEGTNAYSICYFYFREDRETWDADILYC-QNMNSGNLSVLT	OA	EAGFVSLIKES-GT	93		
QY	62	AKIHWIGLRAQNKQKCSIEWDGGSSISKENW-IBESK----	KCLGVH	LETGCFHKWEN	116		
Db	94	DDENVIGLHDPKNNRW--HNSSGLSVYSKSGIGAPSSVNP	GCYSVLS	TSTGTFQWKD	151		
QY	117	FYCEQQDPFVCE	128				
Db	152	VPCEKRFSVCK	163				

Search completed: December 8, 2003, 10:07:06
Job time : 54.0743 secs

inis Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: December 8, 2003, 12:41:24 ; Search time 5475.5 Seconds
(without alignments)
963.811 Million cell updates/sec

Title: US-09-938-114-2
Perfect score: 715
Sequence: 1 DCSGWSSEYGEHCYKVPKQS.....GFKWENFYCQQDPFVCEA 129

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPRO_spool/US09938114/runat_08122003_091000_22849/app_query.fasta_1.725
-DB=genEmbl -QMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09938114 @C9N 1_1_3707 @runat_08122003_091000_22849 -NCPU=6
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb.ba.*
2: gb.hcg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htg.hum.*
40: em.htg.mus.*
41: em.htg.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	701	98.0	733	5	AF176420	AF176420 Deinagkis
2	657	91.9	497	5	AY091758	AY091758 Deinagkis
3	657	91.9	656	5	AB036880	AB036880 Deinagkis
4	653	91.3	544	6	AR259035	AR259035 Sequence
5	496	69.4	678	5	TFLFIXA	D83331 Trimeresuru
6	496	69.4	683	5	AB046491	AB046491 Trimeresuru
7	477	66.7	721	6	AX427204	AX427204 Sequence
8	475	66.4	630	5	AF190827	AF190827 Gloydius
9	460	64.3	632	5	AF125309	AF125309 Gloydius
10	439	61.4	508	5	AF463522	AF463522 Deinagkis
11	439	61.4	661	5	AF540646	AF540646 Deinagkis
12	434	60.7	580	6	AX427207	AX427207 Sequence
13	420.5	58.8	408	5	AF244900	AF244900 Callosela
14	397	55.5	456	6	AX427206	AX427206 Sequence
15	394	55.1	477	5	AY099321	AY099321 Protoboth
16	391	54.7	681	5	AY149341	AY149341 Trimeresuru
17	387	54.1	725	6	AX427210	AX427210 Sequence
18	380.5	53.2	495	5	AY091762	AY091762 Deinagkis
19	378	52.9	432	6	AX427209	AX427209 Sequence
20	377.5	52.8	704	5	AF102901	AF102901 Deinagkis
21	375	52.4	508	5	AY091759	AY091759 Deinagkis
22	375	52.4	666	5	AF540645	AF540645 Deinagkis
23	364	50.9	800	5	CDRNACVXA	Y16348 Crotalus du
24	355.5	49.7	679	5	AB019615	AB019615 Agkistrod
25	329	46.0	474	6	AX427212	AX427212 Sequence
26	291	40.7	698	5	TFLFIXB	D83332 Trimeresuru
27	290	40.6	634	5	AF197915	AF197915 Gloydius
28	290	40.6	688	5	AF540647	AF540647 Deinagkis
29	289	40.4	478	5	AY091761	AY091761 Deinagkis
30	286.5	40.1	690	6	AR036653	AR036653 Sequence
31	286.5	40.1	690	6	AX404807	AX404807 Sequence
32	286.5	40.1	690	6	BD143687	BD143687 Protein h
33	282	39.4	369	5	AY293866	AY293866 Deinagkis
34	282	39.4	574	5	AF350324	AF350324 Deinagkis
35	280	39.2	438	5	AF244901	AF244901 Callosela
36	278	38.9	458	5	AF176421	AF176421 Deinagkis
37	277	38.7	592	5	AB036881	AB036881 Deinagkis
38	274	38.3	478	5	AY091756	AY091756 Deinagkis
39	273	38.2	664	5	AB019616	AB019616 Agkistrod
40	272.5	38.1	583	5	AF125310	AF125310 Gloydius
41	271.5	38.0	487	5	AF463521	AF463521 Deinagkis
42	268.5	37.6	619	5	AF540648	AF540648 Deinagkis
43	264	36.9	483	6	AX427201	AX427201 Sequence
44	264	36.9	738	5	CDRNACVXB	Y16349 Crotalus du
45	256	35.8	484	5	AY091760	AY091760 Deinagkis

ALIGNMENTS

RESULT 1

```

AF176420
LOCUS       AF176420               733 bp    mRNA    linear    VRT 30-APR-2002
DEFINITION   Deinagkistrodon acutus clone 2100490 agkisacutacin A chain mRNA,
complete cds.
ACCESSION   AF176420 AY091757
VERSION     AF176420.2 GI:20273043
KEYWORDS    DEINAGKISTRODON ACUTUS
SOURCE      Deinagkistrodon acutus
ORGANISM    Deinagkistrodon acutus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroides;
            Viperidae; Crotalinae; Deinagkistrodon.
REFERENCE   1 (bases 1 to 733)
AUTHORS    Yu,H.-X., Xiang,K.-J. and Liu,J.
TITLE      cDNA sequencing and analysis of eleven C-type lectin-like protein
            subunits from Agkistrodon acutus
JOURNAL    sheng wu hua hseuh yu sheng wu li hseuh pao (2002) In press
REFERENCE   2 (bases 1 to 733)
AUTHORS    Cheng,X., Qian,Y., Liu,Q., Li,B.X.Y., Ding,J., Xu,Z., Huang,W. and
            Liu,J.
TITLE      Agkisacutacin, a new fibrinolytic & anti-platelet protein from
            Agkistrodon acutus venom
JOURNAL    Unpublished
REFERENCE   3 (bases 1 to 733)
AUTHORS    Yu,H.-X., Xiang,K.-J., Wang,Y. and Liu,J.
TITLE      A chain of agkisacutacin from Deinagkistrodon acutus
JOURNAL    Unpublished
REFERENCE   4 (bases 1 to 733)
AUTHORS    Cheng,X., Liu,J., Li,B.X.Y. and Qian,Y.
TITLE      Direct Submission
JOURNAL    Submitted (08-AUG-1999) Dept. of Biochemistry and Molecular
            Biology, University of Science and Technology of China, School of
            Life-Science, Huangshan Road, Hefei, Anhui 230027, China
REFERENCE   5 (bases 1 to 733)
AUTHORS    Yu,H.-X., Xiang,K.-J., Wang,Y. and Liu,J.
TITLE      Direct Submission
JOURNAL    Submitted (23-APR-2002) Dept. of Biochemistry and Molecular
            Biology, University of Science and Technology of China, School of
            Life-Science, Huangshan Road, Hefei, Anhui 230027, China
REMARK      Sequence update by submitter
COMMENT     On Apr 23, 2002 this sequence version replaced gi:6715112.
FEATURES    Location/Qualifiers
             source
               1..733
               /organism="Deinagkistrodon acutus"
               /mol_type="mRNA"
               /db_xref="taxon:36307"
               /clone="2100490"
               /tissue_type="venom"
               89..547
               /note="consists of two heterologous subunits (A and B
               chains) linked by an inter-subunit disulfide bond; c-type
               lectin family member"
               /codon_start=1
               /product="agkisacutacin A chain"
               /protein_id="AAF26286.2"
               /db_xref="GI:20273044"
               /translation="MGRFTFVSFGLLVFLSLGTAAACSSWSYEGHCYKVFQSK
               TWDAESFCTKVNGGHVLSIESGEADPVGQLIAQKIKSAKHVIGLRAQNKQK
               SIWSDGSSISYENWIEESKKLGVHETGPHKWFYCEQDDPFVCEA"
BASE COUNT   191 a 173 c 204 g 165 t
ORIGIN
Alignment Scores:
Pred. No.:      1,45e-74      Length:      733
Score:          701.00      Matches:    127
Percent Similarity: 98.45%      Conservative: 0
Best Local Similarity: 98.45%      Mismatches: 2
Query Match:    98.04%      Indels:     0
DB:             5          Gaps:      0

US-09-938-114-2 (1-129) x AF176420 (1-733)

QY      1 AspCysSerSerAspTrpSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20

```

```

Db      158 GATTGTTCTCTGTTGGTCTCTCTATGAAGGGCATTTGTTCTTCAACAATCT 217
QY      21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
Db      218 AAGACCTGGACAGATGCAGAGAGCTTCTGCACGAGCAGGTGAACGGGGGCATCTGGTC 277
QY      41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
Db      278 TCTATCGAAGACTCCGAGAACGACACTTTGTGGCCAGTTGATTGCTCAGAGATAAAG 337
QY      61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
Db      338 TCAGCCAAATCCATGCTCGATCGGACTGGAGGCTCAAAACAAAGAAAGCAATGCAGC 397
QY      81 IleGluTrpSerAspGlySerIleSerLysGluAsnTrpIleGluGluGluSerLys 100
Db      398 ATAGAGTGGACGCGATGGCTCCAGCATCATGATTATGAGAAATGGATTGAAGAAGAAATCCAA 457
QY      101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
Db      458 AAGTGTCTTGGGTGCACATAGACAGAGGTTTCATAAGTGGAGAAATTTTACTGTGAA 517
QY      121 GlnGlnAspPropheValCysGluAla 129
Db      518 CAACAAGATCCTTTTGTCTCGAGGCA 544

RESULT 2
AY091758      497 bp    mRNA    linear    VRT 13-MAY-2002
LOCUS       Deinagkistrodon acutus clone 2100590 ACF 1/2 A-chain mRNA, complete
DEFINITION   cds.
ACCESSION   AY091758
VERSION     AY091758.1 GI:20562936
KEYWORDS    DEINAGKISTRODON ACUTUS
SOURCE      Deinagkistrodon acutus
ORGANISM    Deinagkistrodon acutus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroides;
            Viperidae; Crotalinae; Deinagkistrodon.
REFERENCE   1 (bases 1 to 497)
AUTHORS    Yu,H., Xiang,K., Wang,Y. and Liu,J.
TITLE      A chain of ACF 1/2 from Deinagkistrodon acutus
JOURNAL    Unpublished
REFERENCE   2 (bases 1 to 497)
AUTHORS    Yu,H., Xiang,K., Wang,Y. and Liu,J.
TITLE      Direct Submission
JOURNAL    Submitted (22-MAR-2002) Dept. of Biochemistry and Molecular
            Biology, University of Science and Technology of China, School of
            Life Science, Huangshan Road, Hefei, Anhui 230027, P.R. China
FEATURES    Location/Qualifiers
             source
               1..497
               /organism="Deinagkistrodon acutus"
               /mol_type="mRNA"
               /db_xref="taxon:36307"
               /clone="2100590"
               1..459
               /note="C-type lectin family member"
               /codon_start=1
               /product="ACF 1/2 A-chain"
               /protein_id="AAM22786.1"
               /db_xref="GI:20562937"
               /translation="MGRFTFVSFGLLVFLSLGTAAACSSWSYEGHCYKVFQSK
               TWDAESFCTKVNGGHVLSIESGEADPVGQLIAQKIKSAKHVIGLRAQNKQK
               SIWSDGSSISYENWIEESKKLGVHETGPHKWFYCEQDDPFVCEA"
BASE COUNT   135 a 102 c 143 g 117 t
ORIGIN
Alignment Scores:
Pred. No.:      1.68e-69      Length:      497
Score:          657.00      Matches:    119
Percent Similarity: 93.02%      Conservative: 1
Best Local Similarity: 92.25%      Mismatches: 9

```

```

Query Match: 91.89% Indels: 0
DB: 5 Gaps: 0
US-09-938-114-2 (1-129) x AY091758 (1-497)

QY 1 AspCysSerSerAspTyrSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
Db 70 GATTGTCTCTAGTTGGTCTCTCTATGAAGGGCATGTCTACAGGCCTTCAACAATCT 129
QY 21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
Db 130 AAGACCTGGCAGATGTCAGAGAGCTTCTGCACGAAGCAGTGAACGGGGGCATCTGGTTC 189
QY 41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
Db 190 TCTATCGAAGCTCCGGAGAACACACTTTGTGGCCCATCTCTGCTCAGAGATAAAG 249
QY 61 SerAlaLysIleHisValTyrIleGlyLeuAlaGlnAsnLysGlnCysSer 80
Db 250 TCAGCAAAATCCATGCTGGATCGGACTGAGGGCTCAAAACAAGAAAGCAATGCAGC 309
QY 81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluSerLys 100
Db 310 ATAGAGTGGAGCGATGGCTCCAGCATCAGTTATGAGAATTGGATTGAAGAATCCAAA 369
QY 101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
Db 370 AAGTGTCTTGGGGTGCAAAAGCGACAGGGTTTCGTAAGTGGGAGAAATTTTACTGTGAA 429
QY 121 GlnGlnAspPropheValCysGluAla 129
Db 430 CAACGAGATCTTTCGTCGGAGGCA 456

RESULT 3
AB036880
LOCUS Deinagkistrodon acutus acp-a mRNA for anticoagulant protein A, 656 bp mRNA linear VRT 23-MAY-2002
DEFINITION complete cds.
ACCESSION AB036880
VERSION AB036880.1 GI:8980618
KEYWORDS anticoagulant protein A.
SOURCE Deinagkistrodon acutus
ORGANISM Deinagkistrodon acutus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;
Viperidae; Crotalinae; Deinagkistrodon.
REFERENCE 1
AUTHORS Tani, A., Ogawa, T., Nose, T., Nikandrov, N.N., Deshmamaru, M.,
Chijiwa, T., Chang, C.C., Fukumaki, Y. and Ohno, M.
TITLE Characterization, primary structure and molecular evolution of
anticoagulant protein from Agkistrodon acutus venom
JOURNAL Toxicon 40 (6), 803-813 (2002)
MEDLINE 21975857
REFERENCE 2 (bases 1 to 656)
AUTHORS Ogawa, T. and Tani, A.
TITLE Direct Submission
JOURNAL Submitted (06-JAN-2000) Tomohisa Ogawa, Tohoku University, Graduate
School of Agricultural Science, 1-1, Tsutsumidori Amamiyamachi,
Aoba-ku, Sendai 981-8555, Japan (E-mail:ogawa@biochem.tohoku.ac.jp,
URL:http://www.agri.tohoku.ac.jp/hoozo/, Tel:81-22-717-8808,
Fax:81-22-717-8807)
FEATURES
source
1. 656
/organism="Deinagkistrodon acutus"
/mol_type="mRNA"
/db_xref="taxon:36307"
/tissue_type="venom gland"
1. 656
/gene="acp-a"
50. .508
/gene="acp-a"
/codon_start=1
/product="anticoagulant protein A"

sig_peptide
BASE COUNT 173 a 148 c 185 g 150 t
ORIGIN
Alignment Scores:
Pred. No.: 2,36e-69 Length: 656
Score: 657.00 Matches: 119
Percent Similarity: 93.02% Conservative: 1
Best Local Similarity: 92.25% Mismatches: 9
Query Match: 91.89% Indels: 0
DB: 5 Gaps: 0
US-09-938-114-2 (1-129) x AB036880 (1-656)
QY 1 AspCysSerSerAspTyrSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
Db 119 GATTGTCTCTAGTTGGTCTCTCTATGAAGGGCATGTCTATAAGCCCTTCAACAATCT 178
QY 21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
Db 179 AAGACCTGGCAGATGTCAGAGAGCTTCTGCACGAAGCAGGTGAACGGGGGCATCTGGTC 238
QY 41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
Db 239 TCTATCGAAGCTCCGGAGAACGACACTTTGTGGCCCATCTGATTCGTAAGATAAAG 298
QY 61 SerAlaLysIleHisValTyrIleGlyLeuAlaGlnAsnLysGlnCysSer 80
Db 299 TCGCAAAATCCATGCTGGATCGGACTGAGGGCTCAAAACAAGAAAGCAATGCAGC 358
QY 81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluSerLys 100
Db 359 ATAGAGTGGAGCGATGGCTCCAGCATCAGTTATGAGAATTGGATTGAAGAATCCAAA 418
QY 101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
Db 419 AAGTGTCTTGGGGTGCAAAAGCGACAGGGTTTCGTAAGTGGGAGAAATTTTACTGTGAA 478
QY 121 GlnGlnAspPropheValCysGluAla 129
Db 479 CAACGAGATCTTTCGTAAGTGGAGCA 505

RESULT 4
AR259035
LOCUS AR259035
DEFINITION Sequence 1 from patent US 6489451.
ACCESSION AR259035
VERSION AR259035.1 GI:27309520
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 544)
AUTHORS Li, B.X. and Cheng, X.
TITLE Antithrombotic enzyme from the snake venom of agkistrodon acutus
JOURNAL Patent: US 6489451-A 1 03-DEC-2002;
FEATURES
source
1. 544
/organism="unknown"
BASE COUNT 158 a 122 c 143 g 121 t
ORIGIN
Alignment Scores:
Pred. No.: 5.67e-69 Length: 544
Score: 653.00 Matches: 119
Percent Similarity: 99.17% Conservative: 0
Best Local Similarity: 99.17% Mismatches: 1

```



```

gene      1. 683
/ gene="XI/Xbp-a"
CDS       77. 517
/ gene="XI/Xbp-a"
/ codon_start=1
/ product="factor XI/factor X binding protein A chain"
/ protein_id="BAB21452.1"
/ db_xref="GI:12583677"
/ translation="MSFGLLVVAASLRGTGADCLSGWSSYEGHCHYKFEKKTWEDAE
RVCTEQAGHLVSISSGEADFAQLVTQNMKRLDYIWTGLRVQGVKQCNSEWSD
GSSVSYENWIEAESKTCLEKETDPRKWNLYCCQNPVCEA"
BASE COUNT 172 a 162 c 194 g 155 t
ORIGIN

Alignment Scores:
Pred. No.: 4,85e-50 Length: 683
Score: 496.00 Matches: 87
Percent Similarity: 75.97% Conservative: 11
Best Local Similarity: 67.44% Mismatches: 31
Query Match: 69.37% Indels: 0
DB: 5 Gaps: 0

US-09-938-114-2 (1-129) x AB046491 (1-683)
QY 1 AspCysSerSerAspTyrSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
Db 128 GATTGTCTCTGGTGGTCTCTCTATGAGGGGCAATGCTACAGGCCTTCGAAATAAC 187
QY 21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
Db 188 AAGACCTGGGAAGATGTCAGAGAGGCTCTGCAGGAGAGGCGAGCGCGCATCTGGTC 247
QY 41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
Db 248 TCTATCGAAGCTCCGAGAGACAGACTTTGTGGCCCACTGTGTCACACATGAAG 307
QY 61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
Db 308 AGACTCGATTCTATATCTGATCGGGCTGAGGGTTCAAGNAAGGTCAAGCAATGCAC 367
QY 81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluSerLys 100
Db 368 TCGGAGTGGAGCGATGGCTCCAGCGTCAGTTATGAGAACTGGATTGAAGCAGATCCAA 427
QY 101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
Db 428 ACGTGTCTTGGGCTGGAAGAGAGACAGATTTCGTAAGTGGGTCAATATTACTGTGGA 487
QY 121 GlnGlnAspProPheValCysGluAla 129
Db 488 CAACAAATCTCTTCGTCTGCGAGGCA 514

RESULT 7
AX427204 721 bp DNA linear PAT 18-JUN-2002
LOCUS
DEFINITION
Sequence 4 from Patent WO0214364.
ACCESSION
AX427204
VERSION
AX427204.1 GI:21530566
KEYWORDS
Sistrurus miliarius (pigmy rattlesnake)
ORGANISM
Sistrurus miliarius
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;
Viperidae; Crotalinae; Sistrurus.
REFERENCE
1 Sheppard,P.O. and Bishop,P.D.
Rattlesnake venom gland proteins
Patent: WO 0214364-A 4 21-FEB-2002;
ZymoGenetics, Inc. (US)
FEATURES
Location/Qualifiers
1..721
/organism="Sistrurus miliarius"
/mol_type="genomic DNA"

/db_xref="taxon:8758"
91. 549
/ note="ZenK3"
/ codon_start=1
/ protein_id="CAD35499.1"
/ db_xref="GI:21530567"
/ translation="MGREIFVFSGLLVVFLSLSGTGADPCPSGWSYDQHCYRVFKOLK
TWDAERFCSEAGGHLVSISSGEAFVAQLVPENRRALYIWTGLRVQGVKQCNSEWSD
SAKWSGSSVSQYENWIEAESKTCLEQGTNYHKWVNYCGEINPFVCEA"
BASE COUNT 192 a 167 c 203 g 159 t
ORIGIN

Alignment Scores:
Pred. No.: 9,79e-48 Length: 721
Score: 477.00 Matches: 83
Percent Similarity: 76.74% Conservative: 16
Best Local Similarity: 64.34% Mismatches: 30
Query Match: 66.71% Indels: 0
DB: 6 Gaps: 0

US-09-938-114-2 (1-129) x AX427204 (1-721)
QY 1 AspCysSerSerAspTyrSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
Db 160 GATTGTCTCTGGTGGTCTCTCTATGATCAGCATTCACAGGGTCTTCAACAATC 219
QY 21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
Db 220 AAGACGTGGAGCATGTCAGAGAGGTTCTGCTCGGAGCAGCGGAGCGCGCATCTGTC 279
QY 41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
Db 280 TCTATCGAAGCTCCGAGAGACAGCTTTGTGGCCCACTGTGTCACACAGGAGG 339
QY 61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
Db 340 AGACCATCTCTATATCTGATCGACTGAGGGTTCAAGCAAGAGCAATGTCAGC 399
QY 81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluSerLys 100
Db 400 CGGAAGTGGAGCGATGGCTCCAGCGTCAGTTATGAGAACTGGATTGAAGCAGATCCAA 459
QY 101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
Db 460 ACATGTCTTGGGCTGCAACAGGCAACAATATCATAGTGGGTCAATATTACTGTGGA 519
QY 121 GlnGlnAspProPheValCysGluAla 129
Db 520 GAATAAATCTCTTTCGTCTGCGAGGCA 546

RESULT 8
AF190827 630 bp mRNA linear VRT 01-OCT-2000
LOCUS
DEFINITION
Glyodius halyx halyxin A-chain precursor (HXNA) mRNA, complete cds.
ACCESSION
AF190827
VERSION
AF190827.1 GI:10441755
KEYWORDS
Glyodius halyx (halyx viper)
ORGANISM
Glyodius halyx
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;
Viperidae; Crotalinae; Glyodius.
REFERENCE
1 (bases 1 to 630)
Koo,B.H., Sohn,Y.D., Kim,D.S., Jang,Y.S. and Chung,K.H.
A novel coagulation factor Xa inhibitor from Korean snake
(Unpublished)
JOURNAL
2 (bases 1 to 630)
Koo,B.H., Sohn,Y.D., Kim,D.S., Jang,Y.S. and Chung,K.H.
Direct Submission
JOURNAL
Submitted (29-SEP-1999) Cardiovascular Center, Yonsei University
College of Medicine, 134 Shinchon-dong, Seoul 120-752, Republic of
Korea

```

```

FEATURES
  source
    Location/Qualifiers
      1..630
        /organism="Glyodyus halyus"
        /mol_type="mRNA"
        /db_xref="taxon:8714"
      1..630
        /gene="HXNA"
      24..482
        /gene="HXNA"
        /function="coagulation factor Xa inhibitor"
        /codon_start=1
        /product="halyxin A-chain precursor"
        /protein_id="AAG17178.1"
        /db_xref="GI:10441756"
        /translation="WGRFIFVSLGVLVFLSLGTGADPCPSGWSYEGHCYNIFHLFK
        TWAEAFRCRQVGAHLVSISSSEADFAVLSENMKRYGIYIWLIRVGRKKQC
        SSWSDGSSVSQYQNIWIEASKTCLQLQKTEFRKWNFYCGERNPFVCEA"
      24..92
        /gene="HXNA"
      93..479
        /gene="HXNA"
        /product="halyxin A-chain"
      173 a 137 c 174 g 146 t
    ORIGIN
      Alignment Scores:
        Pred. No.: 1,448-47 Length: 630
        Score: 475.00 Matches: 82
        Percent Similarity: 76.74% Conservative: 17
        Best Local Similarity: 63.57% Mismatches: 30
        Query Match: 66.43% Indels: 0
        DB: 5 Gaps: 0
      US-09-938-114-2 (1-129) x AF190827 (1-630)
      QY 1 AspCysSerSerAspTyrSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
      Db 93 GATTGTCCTCTGGTTGGTCTCTCTATGAGGCGATTCACAACTCTTCATCTATC 152
      QY 21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyHisLeuVal 40
      Db 153 AAGACCTGGCAGACAGACAGAGGTTCTGCAGGAAGCAGGTTCGAGGCGCGCATCTGGTTC 212
      QY 41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
      Db 213 TCTATCGAAGTCTCGAAGACAGACACTTTGTGGCCCACTGCTCTCTGAGAACATGAAG 272
      QY 61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGlnLysCysSer 80
      Db 273 AGATACGGGATCTATATCTGGATCGGACTGAGGTTTCGAGGCAAAAGCAATGCAGC 332
      QY 81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluSerLys 100
      Db 333 TCCCACTGGAGCGATGGCTCCAGCTCAGTTATCAGAACTGGATTGAACAGAACTCCAAA 392
      QY 101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
      Db 393 ACCTGTCCTGGCTGCAGGAGAGAGAGAGGTTTCGTAAGTGGTTTAAATTTACTGTGA 452
      QY 121 GlnGlnAspProPheValCysGluAla 129
      Db 453 GAACGAATCCCTTCGTCTGCGAGCA 479
    RESULT 9
    AF125309
    LOCUS
    DEFINITION
      Glyodyus halyus brevicaudus fibrinogen clotting inhibitor A chain
      mRNA, complete cds.
    ACCESSION
      AF125309
    VERSION
      AF125309.1 GI:4337049
    KEYWORDS
      Glyodyus halyus brevicaudus
      Glyodyus halyus brevicaudus
      ORGANISM
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
      Viperidae; Crotalinae; Glyodyus.
    REFERENCE
      1 (bases 1 to 632)
      Kim, D.S. and Koh, Y.S.
      Purification and molecular cloning of snake venom fibrin clotting
      inhibitor
      Unpublished
      2 (bases 1 to 632)
      Kim, D.S. and Koh, Y.S.
      Direct Submission
      JOURNAL
      Submitted (01-FEB-1999) Biochemistry, Yonsei University, Seodaemun
      ku Shinchon dong 134, Seoul 120-749, Korea
    FEATURES
      source
        Location/Qualifiers
          1..632
            /organism="Glyodyus halyus brevicaudus"
            /mol_type="mRNA"
            /sub_species="brevicaudus"
            /db_xref="taxon:66175"
            /note="venom"
          34..498
            /codon_start=1
            /product="fibrinogen clotting inhibitor A chain"
            /protein_id="AAD18055.1"
            /db_xref="GI:4337050"
            /translation="WGRFIFVSLGVLVFLSLGTGADFCPSGWSNNGHCYQAFNQ
            RMTWEDAEFCGSAQAGKGLHLSIETRAEDFAVHVAVRIETSPFHVHGLDEKEQ
            QCSSEWSDGSSVSQYQNIWIEASKTCLGLDELSDNHHKWNVYCGQRNPFVCEA"
          162 a 140 c 182 g 148 t
        ORIGIN
      Alignment Scores:
        Pred. No.: 9,078-46 Length: 632
        Score: 460.00 Matches: 77
        Percent Similarity: 75.78% Conservative: 20
        Best Local Similarity: 60.16% Mismatches: 31
        Query Match: 64.34% Indels: 0
        DB: 5 Gaps: 0
      US-09-938-114-2 (1-129) x AF125309 (1-632)
      QY 2 CysSerSerAspTyrSerTyrGluGlyHisCysTyrLysValPheLysGlnSerLys 21
      Db 112 TGTCCCTCTGGTTGGGCGAGCAATAATGGTCTATCTACAGGCTTCAATCAACGAGT 171
      QY 22 ThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyHisLeuValSer 41
      Db 172 ACCTGGGAGATGCGAGAGAGGTTCTGCTCAGCACAGGCGAGCGGCGCATCTGCTCT 231
      QY 42 IleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLysSer 61
      Db 232 ATCGAAACACGGGCGAGAGCAGACTTTGTGGCCCACTGTCGCTGAGAGGATAGACA 291
      QY 62 AlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGlnLysSerIle 81
      Db 292 TCTTTTCCCATGTCGATCGCACTGAGGATGAAGCAAGCAAGCAAGCAAGCAAGTCT 351
      QY 82 GluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluSerLysLys 101
      Db 352 GAGTGGAGCGATGGATCCAGCGCTCAGTTATGAGAACTGGATTGAAGAGAAATCCAAA 411
      QY 102 CysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGluGln 121
      Db 412 TGTCTTGGCTGGAACACTAGACTCAATATATCATAGTGGTCAATGTTTACTGTGACA 471
      122 GlnAspProPheValCysGluAla 129
      Db 472 AGAAATCCTTTCGTCTGCGAGCT 495
    RESULT 10
    AF463522
    LOCUS
    DEFINITION
      Deinagkistrodon acutus antithrombin 1 A chain mRNA, complete cds.
      508 bp mRNA linear VRT 21-JAN-2002
  
```



```

Db 268 AGTGTGAAGACCATGTCTGACCGGACTGAGGTTCAAAACAAAGAAACAAATGCGAGC 327
Qy 81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluSerLys 100
Db 328 ACAGAGTGGAGGATGGCTCCAGCGTCAGTTATGAGAACCTGCTTGAACATATATGAGA 387
Qy 101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
Db 388 AAGTGTGGGGCGCTGGAACGAGACAGAGGTTTCATAAGTGATGATCAATCTTGGCTGTATA 447
Qy 121 GlnGlnAspPropheValCysGlu 128
Db 448 CAATGAATCTTTCGTCTGCAAG 471

RESULT 12
AX427207
LOCUS AX427207 580 bp DNA linear PAT 18-JUN-2002
DEFINITION Sequence 7 from Patent WO0214364.
ACCESSION AX427207
VERSION AX427207.1 GI:21530569
KEYWORDS
SOURCE
ORGANISM
Sistrurus miliarius (pigmy rattlesnake)
Sistrurus miliarius
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
Viperidae; Crotalinae; Sistrurus.
1
REFERENCE
Sheppard, P.O. and Bishop, P.D.
Rattlesnake venom gland proteins
Patent: WO 0214364-A 7 21-FEB-2002;
ZymoGenetics, Inc. (US)
LOCATION/Qualifiers
1. 580
/organism="Sistrurus miliarius"
/mol_type="genomic DNA"
/db_xref="taxon:8758"
<3. .437
/feature="Zenk4"
/codon_start=1
/protein_id="CAD35500.1"
/db_xref="GI:21530570"
translation="IRNEGTTGADFCPSDWAYDQYCVRVIKQLRTWDDAERFCSEQ
AKGHLVIESGEEAFAVLAQLVAENIKQNDYVIGLRIQSEKQCSIKWSGSSVNY
ENLHKATKKCFGLKXKETGFTWRNVHCTQQLFMCKFPPEC"
BASE COUNT 154 a 133 c 163 g 130 t
ORIGIN

Alignment Scores:
Pred. No.: 1,07e-42 Length: 580
Score: 434.00 Matches: 75
Percent Similarity: 74.22% Conservative: 20
Best Local Similarity: 58.59% Mismatches: 33
Query Match: 60.70% Indels: 0
DB: 6 Gaps: 0

US-09-938-114-2 (1-129) x AX427207 (1-580)

Qy 1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
Db 36 GATTGTCCCTCTGATTGGTATGCTTATGATCACTAGTATTGCTACAGGCTCATCAACAACTC 95
Qy 21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyHisLeuVal 40
Db 96 AGCAGCTGGGACGATCGACGAGAGTTCTCTCGAGCAGCGGAAGCGCGGCATCTCGTC 155
Qy 41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
Db 156 TCTATTGAAGACGACGAGAGCAGCTTGTGTGCCCGCAGCTGCTCGTGAGACATCAAG 215
Qy 61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGlnLysGlnCysSer 80
Db 216 CAAACAAATATGATGTCTGGATCGGACTGAGGATTCAAGCGGCAAGAGACCAATGACG 275

```

```

Qy 81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluSerLys 100
Db 276 ACAGAGTGGAGGATGGCTCCAGCGTCAGTTATGAGAACCTGATTAAACATGCGACCAA 335
Qy 101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
Db 336 AAGTGTGGGGCTGGAAGAAAGACAGAGGTTTCGACGTGGCGCAATGTTCACTGTACA 395
Qy 121 GlnGlnAspPropheValCysGlu 128
Db 396 CAACAAATCTTTCATGTCGAAG 419

RESULT 13
AF244900
LOCUS AF244900 408 bp mRNA linear VRT 29-JUN-2000
DEFINITION Calloselasma rhodostoma aggritin alpha chain mRNA, partial cds.
ACCESSION AF244900
VERSION AF244900.1 GI:9809811
KEYWORDS
SOURCE
ORGANISM
Calloselasma rhodostoma (Malayan pit viper)
Calloselasma rhodostoma
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
Viperidae; Crotalinae; Calloselasma.
1 (bases 1 to 408)
Chung, C.H., Au, L.C. and Huang, T.F.
Molecular cloning and sequence analysis of aggritin, a
collagen-like platelet aggregation inducer
Biochem. Biophys. Res. Commun. 263 (3), 723-727 (1999)
99443731
PUBMED 10512747
REFERENCE
2 (bases 1 to 408)
Chung, C.H., Au, L.C. and Huang, T.F.
Direct Submission
TITLE
Submitted (11-MAR-2000) Pharmacology, College of Medicine, National
Taiwan University, No. 1, Sec. 1, Jen-Ai Rd., Taipei, Taiwan
LOCATION/Qualifiers
1. 408
/organism="Calloselasma rhodostoma"
/mol_type="mRNA"
/db_xref="taxon:8717"
/tissue_type="venom gland"
<1. >408
/feature="Forms heterodimer of alpha and beta chains;
collagen-receptor agonist; induces platelet aggregation;
similar to C-type lectins and glycoprotein Ib binding
proteins"
/codon_start=1
/product="aggritin alpha chain"
/protein_id="AAF79952.1"
/db_xref="GI:8809812"
translation="GLEDCDFGWSYDQHCYQAFNEOKTWDEAEKFCRAQENGAHLAS
IESNGEADFVSWLISOKDELADEYVIGLRAONKEQCSSEWSGSSVSYENLIDLH
TKKCGALEKLGTFRKWNVYCEQMFHAFCKLLPY"
BASE COUNT 115 a 79 c 115 g 99 t
ORIGIN

Alignment Scores:
Pred. No.: 2.87e-41 Length: 408
Score: 420.50 Matches: 77
Percent Similarity: 72.09% Conservative: 16
Best Local Similarity: 59.69% Mismatches: 35
Query Match: 58.81% Indels: 1
DB: 5 Gaps: 1

US-09-938-114-2 (1-129) x AF244900 (1-408)

Qy 1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
Db 10 GATTGTGATTTGGTTGGTCTCCCTATGATCAGCATTTGCTACCAGGATTCATCAATGAACA 69
Qy 21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyHisLeuVal 40

```

```

Db 70 AAAACCTGGATCAGCAGAGAGTTCTGCAGGCCACAGAGAAATGTCGCATCTGGCC 129
QY 41 SerileGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
Db 130 TCTATCGAAAGCAATGGAAGCAGACTTGTTCCTGGCTGATTCTCAGAAAGACGAA 189
QY 61 SerAla---LysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGlnLysGlnCys 79
Db 190 CTGGCAGACGAGACTACCTGCTGGATCGGACTGAGGCTCAAAACAAAGAACAGCAATGC 249
QY 80 SerIleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluSer 99
Db 250 ACCTCGAGTGGAGCGATGCTCCAGTGCAGTATGAGAACTGATTGATCTACATACG 309
QY 100 LysLysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTrpCys 119
Db 310 AAAAAAGTGTGTGCGCTGGAAACAACTGACAGGGTTTCGTAAGTGGCTCAATTACTGT 369
QY 120 GluGlnGlnAspPropheValCysGlu 128
Db 370 GAACAAATGCATGCTTTCGTCTGCAAG 396

```

```

RESULT 14
LOCUS AX427206 456 bp DNA linear PAT 19-JUN-2002
DEFINITION Sequence 6 from Patent WO0214364.
ACCESSION AX427206
VERSION AX427206.1 GI:21530568
KEYWORDS synthetic construct
SOURCE artificial sequences.
ORGANISM Sheppard, P.O. and Bishop, P.D.
REFERENCE Rattlesnake venom gland proteins
AUTHORS Patent: WO 0214364-A 6 21-FEB-2002;
JOURNAL ZymoGenetics, Inc. (US)
FEATURES Location/Qualifiers
source 1..456
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="This degenerate nucleotide sequence encodes the
amino acid sequence of SEQ ID NO:5."
BASE COUNT 72 a 31 c 92 g 66 t 195 others
ORIGIN

```

```

Alignment Scores:
Pred. No.: 2,15e-38 Length: 456
Score: 397.00 Matches: 66
Percent Similarity: 59.6% Conservative: 11
Best Local Similarity: 51.1% Mismatches: 52
Query Match: 55.5% Indels: 0
DB: 6 Gaps: 0

```

US-09-938-114-2 (1-129) x AX427206 (1-456)

```

QY 1 AspCysSerSerAspTrpSerSerTrpGluGlyHisCysTrpLysValPheLysGlnSer 20
Db 70 GAYTGYCCNWSNGGNTGWSNWSNTAYGARCAYTCGTATMGNGTNTTVAARCAVTN 129
QY 21 LysThrTrpTrpAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
Db 130 AARACNTGGAYGAYGCNARGMNTTGTGYWSNGARCARGCGNGNCAYTNGTN 189
QY 41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
Db 190 WSNATHGARWNSNGARGCGNCNTTGTGNCNCARYTNGTNCNCARGAAAYWNGN 249
QY 61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
Db 250 MNGCNCATHYTNTAYATHTGATGTHGGNTNMGNTNCARGGNAARGAARCAATGYWSN 309

```

```

QY 81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluSerLys 100
Db 310 GCNAAARTGWSNGAYGWSNWSNGTNTAYGARAATGGAATGCGARGCGARWNAAR 369
QY 101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTrpCysGlu 120
Db 370 ACNTGYTNGNYTNCARCAGNACNAAYTAYCAAYTGGTNAAYATHTAYTGYGCGN 429
QY 121 GlnGlnAspPropheValCysGluAla 129
Db 430 GARATHAAYCCNTTGTNTGARGCN 456

```

```

RESULT 15
LOCUS AY099321 477 bp mRNA linear VRT 29-MAY-2002
DEFINITION Protobothrops microsquamatus C-type lectin-like protein TMVA A
ACCESSION AY099321
VERSION AY099321.1 GI:21260581
KEYWORDS
SOURCE Protobothrops microsquamatus
ORGANISM Protobothrops microsquamatus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Lepidosauria; Squamata; Scieroglossa; Serpentes; Colubroidea;
Viperidae; Crotalinae; Protobothrops.
REFERENCE 1 (bases 1 to 477)
AUTHORS Wei, Q., Lu, Q.M., Jin, Y., Li, R., Wei, J., Wang, W. and Xiong, Y.
TITLE Purification and cloning of a novel C-type lectin-like protein with
platelet aggregation activity from Trimeresurus microsquamatus
venom
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 477)
AUTHORS Wei, Q., Lu, Q.M., Jin, Y., Li, R., Wei, J., Wang, W. and Xiong, Y.
TITLE Direct Submission
JOURNAL Submitted (23-APR-2002) Department of Animal Toxicology, Kunming
Institute of Zoology, 32 East Jiaochang Road, Kunming, Yunnan
650223, China
FEATURES Location/Qualifiers
source 1..477
/organism="Protobothrops microsquamatus"
/mol_type="mRNA"
/isolation_source="venom"
/db_xref="taxon:103944"
CDS 1..477
/note="platelet aggregation activator"
/codon_start=1
/product="C-type lectin-like protein TMVA A chain"
/protein_id="AA043808.1"
/db_xref="GI:21260582"
/translation="MGRFTFVSFGLVFLVLSGTFADFCIPGWSAYDRVCYQAFSE
PKNWEDAESFCIEGVKTSHLVSISSGEGDFVAOLVAEKIKTSFYVWIGRIQNKEQ
QCRSEWSDASSVNYENLFKQSSKKCYALKKGTGLRTWNVYCGRENPFVKYTEC"
BASE COUNT 122 a 100 c 134 g 121 t
ORIGIN

```

Alignment Scores:

```

Pred. No.: 5,2e-38 Length: 477
Score: 394.00 Matches: 69
Percent Similarity: 70.31% Conservative: 21
Best Local Similarity: 53.91% Mismatches: 38
Query Match: 55.10% Indels: 0
DB: 5 Gaps: 0

```

US-09-938-114-2 (1-129) x AY099321 (1-477)

```

QY 1 AspCysSerSerAspTrpSerSerTrpGluGlyHisCysTrpLysValPheLysGlnSer 20
Db 76 GATTGTATCCCTGGTGTGTCGCGCTATGATCGGTATGTTACCAAGGCTTCAGTGAACCG 135
QY 21 LysThrTrpTrpAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
Db 136 AAAAACTGGGAAGATGACAGAGTTTCTGCGAGGAGGGGTGAAGACCTCGCATCTGGTC 195

```

```
QY 41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
196 TCCATCGAAAGCTCCGGAGAGGAGACTTCGTGGCCAGCTGGTCGCTGAGAAGATAAG 255
QY 61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
Db ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
256 ACGTCCTTTACGATGTCTGGATTGGCTGGCTGAGGATTCAAAACAAAAGACAGCAATGCAGG 315
QY 81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluSerLys 100
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
316 TCGGAGTGGAGCGATGCCCTCCAGTGTCAATTATGAGAACTTGTTTAAACAATCTTCCAAA 375
QY 101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
Db ||||| ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
376 AAATGTTATGCCCTGAAAAAGGACAGAGCTTCGCACGTGGTTCAATGTTTACTGTGGA 435
QY 121 GlnGlnAspProPheValCysGlu 128
Db ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
436 AGAGAAATTCCTTTCGTCTGCAAG 459
```

Search completed: December 8, 2003, 16:08:44
Job time : 5477.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 8, 2003, 12:39:59 ; Search time 443.023 Seconds
(without alignments)
786.026 Million cell updates/sec

Title: US-09-938-114-2

Perfect score: 715

Sequence: 1 DCSSDWSSYEGHCYKVKQS.....GFKWENYCFQDPFFVCEA 129

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlh
-Q=/cgn2.1/USPTO_spool/US09938114/runat_08122003_091000_22839/app query.fasta_1.725
-DB=N_Geneseq_19Jun03 -QMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09938114 -CGN 1 1 0 @runat_08122003_091000_22839 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCAPOP=6
-FCAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_19Jun03.*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	653	91.3	544	25	ABX93668	cDNA encoding Dein
2	477	66.7	721	24	AAD32055	Pigmy rattlesnake
3	475	66.4	601	22	AAI71876	Snake venom blood
4	460	64.3	632	24	AAL42015	Korean adder snake
5	434	60.7	580	24	AAD32057	Pigmy rattlesnake
6	397	55.5	456	24	AAD32056	Pigmy rattlesnake
7	387	54.1	725	24	AAD32059	Pigmy rattlesnake
8	378	52.9	432	24	AAD32058	Pigmy rattlesnake
9	329	46.0	474	24	AAD32060	Pigmy rattlesnake
10	290	40.6	633	22	AAI71877	Snake venom blood
11	286.5	40.1	690	16	AAQ89309	Snake venom antith
12	286.5	40.1	690	21	AAC61144	DNA encoding a sna
13	286.5	40.1	690	24	AAK99834	DNA encoding the a
14	272.5	38.1	583	24	AAL42016	Korean adder snake
15	264	36.9	483	24	AAD32053	Pigmy rattlesnake
16	237.5	33.2	716	17	AAT64829	Snake venom protei
17	216	30.2	454	25	ABX93674	cDNA encoding Dein
18	210	29.4	453	24	AAD32054	Pigmy rattlesnake
19	201.5	28.2	441	11	AAQ05622	Sequence encoding
20	201.5	28.2	562	22	AAS31360	Human cDNA encodi
21	201.5	28.2	562	24	ABQ66684	Human polynucleoti
22	201.5	28.2	777	21	AAZ99820	cDNA sequence of h
23	201.5	28.2	800	22	AAZ75077	Human colon associ
24	201.5	28.2	843	21	AAZ75077	Human cancer associ
25	201.5	28.2	843	21	AAZ75090	Human colon associ
26	201.5	28.2	1066	22	AAZ31201	Human cDNA encodi
27	201.5	28.2	1066	24	ABQ66525	Human polynucleoti
28	201.5	28.2	4771	18	AAV02186	Human type C lecti
29	201.5	28.2	4771	21	AAZ92447	Human type C lecti
30	201.5	28.2	5641	24	ABL92107	Human Tumour Endot
31	201.5	28.2	5641	25	ABX72032	DNA encoding human
32	201.5	28.2	5641	25	ABZ34758	Coding sequence SE
33	201.5	28.2	5747	21	AAZ7360	Human ORF2915
34	198.5	27.8	678	24	ABZ11665	Human polynucleoti
35	197.5	27.6	498	9	AAZ81962	Sequence of human
36	197.5	27.6	4451	25	ABZ34683	Coding sequence SE
37	197.5	27.6	4588	18	AAV02185	Murine type C lect
38	197.5	27.6	4588	21	AAA32446	Murine E-selectin
39	197.5	27.6	4588	24	ABL92140	Mouse Tumour Endot
40	197.5	27.6	4588	25	ABX72065	DNA encoding mouse
41	197.5	27.6	5837	21	AAA92454	Murine E-selectin
42	196.5	27.5	495	14	AAZ40477	MURF-1 coding reg
43	195.5	27.3	495	10	AAZ91109	Human reg cDNA, H
C 44	193.5	27.1	710	24	ABV95090	Human pancreatic c
C 45	192.5	26.9	539	24	ABV98260	Human pancreatic c

ALIGNMENTS

RESULT 1

ABX93668
ID ABX93668 standard; cDNA; 544 BP.

AC ABX93668;

DT 02-JUN-2003 (first entry)

XX cDNA encoding Deinagkistrodon acutus antithrombosis enzyme alpha chain.

XX Antithrombosis; ss; PCR; alpha chain; fibrin hydrolysis; blood clot;
XX platelet aggregation; vaso-occlusive disorder; thromboembolic disorder;
XX myocardial infarction; restenosis; cancer; neurodegenerative disease;
XX angioathic thrombosis; cerebral thrombosis; thromboangiitis obliterans;
XX ischaemic cerebral vascular disease; unstable angina; acute thrombosis;
XX unstable stenocardia; pulmonary embolism; deep vein thrombosis; oedema;
XX peripheral arterial occlusion; stroke; atherosclerosis; inflammation;

thrombosis.

Deinagkistrodon acutus.

Key Location/Qualifiers

CDS 1..363

FT /*tag= a

FT /product= "Antithrombosis enzyme alpha chain"

FT /partial

FT /transl_except= (pos:244..246,aa:Lys)

FT /note= "No start codon given"

US6489451-B1.

03-DEC-2002.

10-APR-1998; 98US-0058740.

10-APR-1997; 97US-043886P.

(HEFE-) HEFEI SIU FUNG USTC PHARM CO LTD.

Li BX, Cheng X;

WPI; 2003-352116/33.

P-PSDB; AB008798.

New purified Agkistrodon actus anti-thrombosis enzyme, useful for preventing and treating vaso-occlusive and thromboembolic disorders, including myocardial infarction, restenosis, cerebral thrombosis and unstable angina

Example 3; Column 25-26; 19pp; English.

The invention relates to a new Deinagkistrodon actus anti-thrombosis enzyme, where the enzyme hydrolyses fibrin, dissolves blood clots and prevents platelet aggregation. The anti-thrombosis enzyme was administered to rabbits intravenously. Thrombosis was determined before and after the administration at 0.5, 1.0, 2.0 and 3.0 hours by the Chandler method. The enzyme showed anti-thrombosis activity at 0.5 hour following administration at 0.005 micro/kg and this activity was increased significantly at 1.0 hour and at 0.01 micro /kg. The composition is useful for preventing and treating vaso-occlusive and thromboembolic disorders, including myocardial infarction, restenosis, angiothritic thrombosis, cerebral thrombosis, ischaemic cerebral vascular diseases, unstable angina, acute thrombosis, unstable stenocardia, thromboangitis obliterans, pulmonary embolism, deep vein thrombosis, peripheral arterial occlusion, stroke. It is also useful for treating atherosclerosis, oedema and inflammation, cancer and neurodegenerative diseases. The present sequence represents cDNA encoding the Deinagkistrodon actus antithrombosis enzyme alpha chain.

Sequence 544 BP; 158 A; 122 C; 143 G; 121 T; 0 other;

Alignment Scores:

Pred. No.: 5,36e-69 Length: 544

Score: 653.00 Matches: 119

Percent Similarity: 99.17% Conservatives: 0

Best Local Similarity: 99.17% Mismatches: 1

Query Match: 91.33% Indels: 0

DB: 25 Gaps: 0

US-09-938-114-2 (1-129) x ABX93668 (1-544)

QY 10 GluGlyHisCysTyrLysValPheLysGlnSerLysThrTrpThrAspAlaGluSerPhe 29

DB 1 GAAGGGCAATGTGCTCAAGGCTTCAACCAATCTAAGACCTGACAGATCCAGAGAGCTTC 60

QY 30 CysThrLysGlnValAsnGlyGlyHisLeuValSerIleGluSerSerGlyGluAlaAsp 49

DB 61 TCACAGACAGGTGAACGGGGGCATCTGGTCTTATCGAAGCTCCGAGAGACAGAC 120

QY 50 PheValGlyGlnLeuIleAlaGlnLysIleLysSerAlaLysIleHisValTrpIleGly 69

Db 121 TTTGTGGCCAGTTGATGTCTCAGAGATAGTCAAGCAAAATCCATGTCTGGATCGA 180

QY 70 LeuArgAlaGlnAsnLysGlnCysSerIleGluTrpSerAspGlySerSerIle 89

Db 181 CTGAGGGCTCAAAACAAAGAAAGCAATGCAGCATAGAGTGGGATGGCTCCAGCATC 240

QY 90 SerLysGluAsnTrpIleGluGluSerLysLysCysLeuGlyValHisIleGluThr 109

Db 241 AGTTATGAGAATTGGATTGAAGAAATCCAAAGAGTGTCTTGGGTGCACATAGAGACA 300

QY 110 GlyPheHisLysTrpGluAsnPheTyrCysGluGlnGlnAspPropheValCysGluAla 129

Db 301 GGGTTTCATAAGTGGAGAAATTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAGCA 360

RESULT 2

AA032055

ID AAD32055 standard; DNA; 721 BP.

XX

AC AAD32055;

XX

DT 18-JUN-2002 (first entry)

XX

DE Pigmy rattlesnake venom gland protein, Zsnk3 gene.

XX

DE Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;

XX

KW platelet aggregation; gene; Zsnk3; ds.

XX

OS Sistrurus miliarius.

XX

PH Key

CDS Location/Qualifiers

FT 91..549

FT /*tag= a

FT /product= "Zsnk3 protein"

FT sig_peptide 91..159

FT /*tag= b

FT mat_peptide 160..546

FT /*tag= c

FT /product= "Mature Zsnk3 protein"

XX

PN WO200214364-A2.

XX

PD 21-FEB-2002.

XX

PD 13-AUG-2001; 2001WO-US25310.

XX

PR 14-AUG-2000; 2000US-225072P.

XX

PR 14-AUG-2000; 2000US-225087P.

XX

PR 15-AUG-2000; 2000US-225489P.

XX

PR 15-AUG-2000; 2000US-225490P.

XX

PR 20-DEC-2000; 2000US-356997P.

XX

XX (ZYMO) ZYMOGENETICS INC.

XX

PI Sheppard PO, Bishop PD;

XX

XX WPI; 2002-269180/31.

XX

DR P-PSDB; AAE20179.

XX

PT New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which affect blood coagulation and platelet aggregation system, useful in therapy and diagnostics, or as tools in the study of genetics or molecular biology

XX

PS Claim 5; Page 73-74; 79pp; English.

XX

CC The invention relates to new pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which affect blood coagulation and platelet aggregation system. The polypeptides, which affect blood coagulation and platelet aggregation system, are useful in therapy and diagnostics. The polypeptides are also useful as an educational tool in laboratory practical kits for courses related to genetics and molecular biology, protein chemistry and antibody production and analysis. The

FH Key Location/Qualifiers
 FT CDS 34..498
 FT /tag= a
 FT /product= "Salmorin A chain protein"
 FT sig_peptide 34..102
 FT /tag= b
 FT /note= "Signal peptide"
 FT mat_peptide 103..495
 FT /tag= c
 FT /note= "Mature Salmorin A chain protein"
 FT 3'UTR 499..622
 FT /tag= d
 FT polyA_signal 623..632
 FT /tag= e
 FT
 XX WO200214514-A1.
 PN
 XX
 XX 21-FEB-2002.
 XX
 XX 26-JUL-2001; 2001WO-KR01277.
 XX
 XX 27-JUL-2000; 2000KR-0043470.
 XX (BIOB-) BIOBUD CO LTD.
 XX Chung K, Kim D, Koh Y;
 XX WPI; 2002-241907/29.
 DR P-PSDB; AA014520.
 DR
 XX New salmorin protein derived from venom of Korean adder Agkistrodon
 PT halya brevicaudus, useful for treating thrombosis by repressing
 PT fibrinogen clotting through repression of activation of prothrombin
 PT into thrombin
 XX
 XX Claim 1; Fig 1A; 30pp; English.
 XX
 CC The invention comprises the nucleotide and protein sequences of a
 CC salmorin protein derived from the venom of Korean adder. Salmorin protein
 CC is composed of an A chain and a B chain, and has inhibitory activity
 CC against fibrinogen clotting. The salmorin protein of the invention is
 CC useful for treating thrombosis, as it represses fibrinogen clotting
 CC potentially by binding to prothrombin and thrombin so as to delay blood
 CC coagulation. The present cDNA sequence encodes the Korean adder salmorin
 CC A chain protein.
 XX
 SQ Sequence 632 BP; 162 A; 140 C; 182 G; 148 T; 0 other;
 Alignment Scores:
 Pred. No.: 9 53e-46 Length: 632
 Score: 460.00 Matches: 77
 Percent Similarity: 75.78% Conservative: 20
 Best Local Similarity: 60.16% Mismatches: 31
 Query Match: 64.34% Indels: 0
 DB: 24 Gaps: 0
 US-09-938-114-2 (1-129) x AAL42015 (1-632)
 QY 2 CysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSerLys 21
 Db 112 TGTCCCTCTGTTGGGCGAGCAATATGCTATGCTACGAGCCCTTCAATCAACGGATG 171
 QY 22 ThrTrpTrpAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuValSer 41
 Db 172 ACTGGGAGATGCGACAGAGTTCTCTCAGCACAGCGGCGGCGATCTGGTCTCT 231
 QY 42 IleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLysSer 61
 Db 232 ATCGAAACAGCGGCAGACAGACAGACTTTGTGGCCACGCTGCTCAGAGGATAGAGACA 291
 QY 62 AlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGlnLysGlnCysSerIle 81
 Db 292 TCTTTTCCCATCTCTGGTGGACTGAGGGATGAAGGCAAGAACAGCAATCGAGTCT 351

QY 82 GluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluSerLysLys 101
 Db 352 GAGTGGAGCGATGGATCCAGCGTCAGTTATGAGAACTGGATTGAAGCAGAAATCCAAACA 411
 QY 102 CysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGluGln 121
 Db 412 TGTCTTGGCTGGAACTAGACTCAAAATATCATAGTGGTCAATGTTTACTGTGGACAA 471
 QY 122 GlnAspPropheValCysGluAla 129
 Db 472 AGAAATCCTTTCTGCTCGAGGCT 495
 RESULT 5
 AAD32057
 ID AAD32057 standard; DNA; 580 BP.
 XX AC
 XX AAD32057;
 XX 18-JUN-2002 (first entry)
 XX DE Pigmy rattlesnake venom gland protein, Zsnk4 gene.
 XX KW Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;
 XX platelet aggregation; gene; Zsnk4; ds.
 XX OS Sistrurus miliarius.
 XX FH Key Location/Qualifiers
 FT CDS 3..437
 FT /tag= a
 FT /product= "Zsnk4 protein"
 FT /note= "CDS does not include start codon"
 FT /partial
 FT sig_peptide 3..29
 FT /tag= b
 FT mat_peptide 30..434
 FT /tag= c
 FT /product= "Mature Zsnk4 protein"
 FT
 XX WO200214364-A2.
 XX 21-FEB-2002.
 XX 13-AUG-2001; 2001WO-US25310.
 XX 14-AUG-2000; 2000US-225072P.
 XX 14-AUG-2000; 2000US-225087P.
 XX 15-AUG-2000; 2000US-225489P.
 XX 15-AUG-2000; 2000US-225490P.
 XX 20-DEC-2000; 2000US-356997P.
 XX (ZYMO) ZYMOGENETICS INC.
 XX Sheppard PO, Bishop PD;
 XX WPI; 2002-269180/31.
 DR P-PSDB; AAE20180.
 XX New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which
 PT affect blood coagulation and platelet aggregation system, useful in
 PT therapy and diagnostics, or as tools in the study of genetics or
 PT molecular biology
 XX Claim 5; Page 75; 79pp; English.
 XX The invention relates to new pigmy rattlesnake (Sistrurus miliarius)
 CC venom gland proteins, which affect blood coagulation and platelet
 CC aggregation system. The polypeptides, which affect blood coagulation and
 CC platelet aggregation system, are useful in therapy and diagnostics. The
 CC polypeptides are also useful as an educational tool in laboratory
 CC practical kits for courses related to genetics and molecular biology,
 CC protein chemistry and antibody production and analysis. The

CC polynucleotide or polypeptide can be used as standards or as unknowns
CC for testing purposes. The polypeptides are also useful in identifying
CC proteins by western blotting, protein purification, determining the
CC weight of expressed polypeptides as a ratio to total protein expressed,
CC identifying peptide cleavage sites, coupling amino and carboxyl terminal
CC tags, mass spectrometry, circular dichroism to determine conformation or
CC affinity chromatography columns to purify the protein, cloning or
CC sequencing. The present sequence is *Sistrurus miliarius* venom gland
CC protein, Zsnnk4 gene.

SQ Sequence 580 BP; 154 A; 133 C; 163 G; 130 T; 0 other;

Alignment Scores:		
Pred. No.:	1.12e-42	Length:
Score:	434.00	Matches:
Percent Similarity:	74.22%	Conservative:
Best Local Similarity:	58.59%	Mismatches:
Query Match:	60.70%	Indels:
DB:	24	Gaps:
		0
		580

US-09-938-114-2 (1-129) x AAD32057 (1-580)

QY	1	AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrIysValPheLysGlnSer	20
Db	36	GATTGTCCCTCTGATTGGTATGCCCTATCATCAGGGTCATCAACAACATC	95
QY	21	LysThrTrpThrAspAlaGluSerPheCysThrIysGlnValAsnGlyGlyHisLeuVal	40
Db	96	AGGACGTGGAGCATGCAGAGAGGTTCTGTCGAGCAGCGAAGCGCGGCATCTCGTC	155
QY	41	SerIleGluSerSerGlyGluAlaaspPheValGlyGlnLeuIleAlaGlnLysIleLys	60
Db	156	TCTATTGAAGCGACGGAGACGCCCTTTGTGCCCCAGCTGGTCGTGAGAACATCAAG	215
QY	61	SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer	80
Db	216	CAAAAACAAATATGATGTCTCGATCGACTGAGGATTCAAGCGCAGAGCAATCGAC	275
QY	81	IleGluTrpSerAspGlySerSerIysSerLysGluAsnTrpIleGluGluSerLys	100
Db	276	ACGAAGTGGAGCGATGGCTCCAGCGCTCAATTATGAGAACCCTGATTAAACATGCGACCAA	335
QY	101	LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu	120
Db	336	AGTGTGTTGGCTCAAAAAGAGACAGGGTTTCGCACGTGGCGCAATGTTCACTGTACA	395
QY	121	GlnGlnAspProPheValCysGlu	128
Db	396	CAACAAAATCTTTTCATGTGCAAG	419

RESULT 6

AAD32056

ID AAD32056 standard; DNA; 456 BP.

[illegible]

AC AAD32056;

XX

DT 18-JUN-2002 (first entry)

XX

DE Pigmy rattlesnake venom gland protein Zsnk3, degenerate nucleic acid.

XX Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;
KW platelet aggregation; ds.
KW

OS *Sistrurus miliarius*.

XX PN WO200214364-A2.

XX
PD 21-FEB-2002.

XX
PF 13-AUG-2001; 2001WO-US25310.
XX

XX
PR 14-AUG-2000; 2000US-225072P.

PR 14-AUG-2000; 2000US-225087P.

PR 15-AUG-2000; 2000US-225489P.
PR 15-AUG-2000; 2000US-225490P.
PR 20-DEC-2000; 2000US-356997P.

PA (ZYMO) ZYMOGENETICS INC.

XX
PI Sheppard PO, Bishop PD;
XX

XX WPI; 2002-269180/31.

DR P-PSDB; AAE20179.

XX

PT New pigmy rattlesnake (*Sistrurus miliarius*) venom gland proteins, which
PT affect blood coagulation and platelet aggregation system, useful in
PT therapy and diagnostics, or as tools in the study of genetics or
PT molecular biology

PS Disclosure; Page 74; 79pp; English.

The invention relates to new pigmy rattlesnake (*Sistrurus miliarius*) venom gland proteins, which affect blood coagulation and platelet aggregation system. The polypeptides, which affect blood coagulation and platelet aggregation system, are useful in therapy and diagnostics. The polypeptides are also useful as an educational tool in laboratory practical kits for courses related to genetics and molecular biology, protein chemistry and antibody production and analysis. The polynucleotide or polypeptide can be used as standards or as unknowns for testing purposes. The polypeptides are also useful in identifying proteins by western blotting, protein purification, determining the weight of expressed polypeptides as a ratio to total protein expressed, identifying peptide cleavage sites, coupling amino and carboxyl terminal tags, mass spectrometry, circular dichroism to determine conformation or affinity chromatography columns to purify the protein, cloning or sequencing. The present sequence is *Sistrurus miliarius* venom gland protein Zsnk3, degenerate nucleic acid.

SQ Sequence 456 BP; 72 A; 31 C; 92 G; 66 T; 195 other;

Alignment Scores:

Pred. No.:	2,23e-38	Length:	456
Score:	397.00	Matches:	66
Percent Similarity:	59.69%	Conservative:	11
Best Local Similarity:	51.16%	Mismatches:	52
Query Match:	55.52%	Indels:	0
DB:	24	Gaps:	0

US-09-938-114-2 (1-129) x AAD32056 (1-456)

QY 1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
 ||||| ||| |||::| ||||| ||||| ||||| |||||
 Db 70 GAYTGYCCNWSNGGNTGGWSNWSNTAYGAYCARCAYTGYTAYMGNGTNTTVAARCARYTN 129

Qy 21 LysThrTrrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
||||| ||||| ||||| ||||| :||| ||||| ||||| |||||
Db 130 AARACNTGGGAYGAYGCNGARMNNTYTYGSNGARCARGCGNGGNCAYTYNGTN 189

Qy 41 SerIleGluSerSerGlyClnAlaAspPheValGlyGlnLeuIlealaGlnLysIleLys 60
 ||||| ||||| ||||| ||||| |||||
Db 190 WSNATHGRWNSWNWGARGGCNCNTTYGTNGCNCARYTNGTCNCGARAAAYMGNGN 249

Qy 61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
 ||| ::: ::| | | | | | | | | |
Dδ 250 MGNGCNATHYNTAYATHTGGATHGNNYTMNGTNCARGGNAAAGARACARTCYWSN 309

Qy 81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluSerLys 100
:::||||| ::::|
Db 310 GCGAARTCGWNGAYGGWNSWNGTNNWNTAYGARAAYTGATCGCGCGARNSAAR 369

Qy 101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
 |||||
 ||||| : : ||||| |||||
 Db 370 A NTGTYNGGNTNCAPCARGNACNAAYTAYCAYAAATGGGTAAAYATHATYATGYGNN 429

121 GlnGlnAspPropheValCysGluAla 129 QY

1			
1			
3			
1			
9			
9			
1			
3			
0			
0			
1			
0			
1			
2			
4
0			
0			
1	
1	
1	
1			
1			
2			
2			
2			

Db 430 GARATHAAYCCTTGTGTCGACGCN 456

RESULT 7

AAD32059

ID AAD32059 standard; DNA; 725 BP.

XX

AC AAD32059;

XX

XX 18-JUN-2002 (first entry)

XX Pigmy rattlesnake venom gland protein, Zsnk5 gene.

DE

XX

XX Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;

KW platelet aggregation; gene; Zsnk5; ds.

XX

OS Sistrurus miliarius.

XX

XX Key Location/Qualifiers

PH 88..564

FT CDS

FT /tag= a

FT /product= "Zsnk5 protein"

FT sig_peptide

FT 88..156

FT /tag= b

FT mat_peptide

FT 157..561

FT /tag= c

FT /product= "Mature Zsnk5 protein"

XX

XX WO200214364-A2.

XX

XX 21-FEB-2002.

XX

XX 13-AUG-2001; 2001WO-US25310.

XX

XX 14-AUG-2000; 2000US-225072P.

PR 14-AUG-2000; 2000US-225087P.

PR 15-AUG-2000; 2000US-225489P.

PR 15-AUG-2000; 2000US-225490P.

PR 20-DEC-2000; 2000US-356997P.

XX

XX (ZYMO) ZYMOGENETICS INC.

XX

XX Sheppard PO, Bishop PD;

XX

XX WPI; 2002-269180/31.

DR P-PSDB; AAE20181.

XX

XX New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which

PT affect blood coagulation and platelet aggregation system, useful in

PT therapy and diagnostics, or as tools in the study of genetics or

PT molecular biology

XX

XX Claim 5; Page 77; 79pp; English.

XX

XX The invention relates to new pigmy rattlesnake (Sistrurus miliarius)

CC venom gland proteins, which affect blood coagulation and platelet

CC aggregation system. The polypeptides, which affect blood coagulation and

CC platelet aggregation system, are useful in therapy and diagnostics. The

CC polypeptides are also useful as an educational tool in laboratory

CC practical kits for courses related to genetics and molecular biology,

CC protein chemistry and antibody production and analysis. The

CC polynucleotide or polypeptide can be used as standards or as unknowns

CC for testing purposes. The polypeptides are also useful in identifying

CC proteins by western blotting, protein purification, determining the

CC weight of expressed polypeptides as a ratio to total protein expressed,

CC identifying peptide cleavage sites, coupling amino and carboxyl terminal

CC tags, mass spectrometry, circular dichroism to determine conformation or

CC affinity chromatography columns to purify the protein, cloning or

CC sequencing. The present sequence is Sistrurus miliarius venom gland

CC protein, Zsnk5 gene.

XX

XX Sequence 725 BP; 179 A; 171 C; 200 G; 175 T; 0 other;

SQ

Alignment Scores:

Pred. No.: 6.62e-37 Length: 725

Score: 387.00 Matches: 67

Percent Similarity: 71.88% Conservative: 25

Best Local Similarity: 52.34% Mismatches: 36

Query Match: 54.13% Indels: 0

DB: 24 Gaps: 0

US-09-938-114-2 (1-129) x AAD32059 (1-725)

QY 1 AspCysSerSerAspTyrGluGlyHisCysTyrLysValPheLysGlnSer 20

DB 163 AATTGTCCTCTGGTTGGTTCGCTACGATCAGTATTGCTACAGGTCATCAACGACTC 222

QY 21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40

DB 223 AAGACCTGGGACGATGCGAGCGGTTCTGCTGGAGCAGCGGAGGCGGGCATCTGGCG 282

QY 41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60

DB 283 TCTGTCGAAACGATGAAGACGACTCTTCTGGCCAGTTGGTCTCGCAACATAAG 342

QY 61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGlnLysCysSer 80

DB 343 CAAAACCAATATCTATGTTGGATTGCTGAGGATTCAAAACAAAGGACAGCAATGCAGC 402

QY 81 IleGluTrpSerAspGlySerSerIleSerLysGluAenTrpIleGluGluSerLys 100

DB 403 ACGAAGTGGAGCGATGGCTCCAGCTCAGTTATGAGAACCTGGTTAAATCACAATCCAAA 462

QY 101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120

DB 463 AAGTGTGTTGGCTGAAAAAGAGACAGAGATTCTTCAATGATGATCAATGCTGCTGCGAA 522

QY 121 GlnGlnAspPropheValCysGlu 128

DB 523 GAAAAAACCTTTCTGCTGCAAG 546

RESULT 8

AAD32058

ID AAD32058 standard; DNA; 432 BP.

XX

AC AAD32058;

XX

XX 18-JUN-2002 (first entry)

DT

XX

XX Pigmy rattlesnake venom gland protein Zsnk4, degenerate nucleic acid.

DE

XX

XX Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;

KW platelet aggregation; ds.

XX

XX Sistrurus miliarius.

XX

XX WO200214364-A2.

XX

XX 21-FEB-2002.

XX

XX 13-AUG-2001; 2001WO-US25310.

XX

XX 14-AUG-2000; 2000US-225072P.

PR 14-AUG-2000; 2000US-225087P.

PR 15-AUG-2000; 2000US-225489P.

PR 15-AUG-2000; 2000US-225490P.

PR 20-DEC-2000; 2000US-356997P.

XX

XX (ZYMO) ZYMOGENETICS INC.

XX

XX Sheppard PO, Bishop PD;

XX

XX WPI; 2002-269180/31.

DR

XX

XX New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which

PT affect blood coagulation and platelet aggregation system, useful in

PT therapy and diagnostics, or as tools in the study of genetics or

PT

PT molecular biology -
 XX Disclosure; Page 76; 79pp; English.
 XX The invention relates to new pigmy rattlesnake (Sistrurus miliarius)
 CC venom gland proteins, which affect blood coagulation and platelet
 CC aggregation system. The polypeptides, which affect blood coagulation and
 CC platelet aggregation system, are useful in therapy and diagnostics. The
 CC polypeptides are also useful as an educational tool in laboratory
 CC practical kits for courses related to genetics and molecular biology,
 CC protein chemistry and antibody production and analysis. The
 CC polynucleotide or polypeptide can be used as standards or as unknowns
 CC for testing purposes. The polypeptides are also useful in identifying
 CC proteins by western blotting, protein purification, determining the
 CC weight of expressed polypeptides as a ratio to total protein expressed,
 CC identifying peptide cleavage sites, coupling amino and carboxyl terminal
 CC tags, mass spectrometry, circular dichroism to determine conformation or
 CC affinity chromatography columns to purify the protein, cloning or
 CC sequencing. The present sequence is Sistrurus miliarius venom gland
 CC protein Zank4, degenerate nucleic acid.
 XX
 SQ Sequence 432 BP; 91 A; 34 C; 84 G; 55 T; 168 other;
 Alignment Scores:
 Pred. No.: 3.96e-36 Length: 432
 Score: 378.00 Matches: 64
 Percent Similarity: 62.50% Conservative: 16
 Best Local Similarity: 50.00% Mismatches: 48
 Query Match: 52.87% Indels: 0
 DB: 24 Gaps: 0
 US-09-938-114-2 (1-129) x AAD32058 (1-432)
 Qy 1 AspCysSerSerAspTrpSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
 Db 34 GAYTGCCNWSNGAYTGGTAYGNCATYAYGAYCARTAYTGYTMGMNTNATHAARCARVTN 93
 Qy 21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
 Db 94 MGNACNTGGGAYGCGNGARGMNTTYTGYMSNGARGCARGCNAARGGNGNCAYTYNTGN 153
 Qy 41 SerLeuGluSerSerGlyGluAlaAspPheValGlyGlnLeuLeuAlaGlnLysIleLys 60
 Db 154 WSNATHGARWSNGAYGNGARGCNCNTTYTGNCNCARYNTGNCNGARAAYATHAAR 213
 Qy 61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
 Db 214 CARAAYAAATAYAGTNTGGATGNGNTNMGNATHCARGGNGARGAARACARTGYWSN 273
 Qy 81 IleGluTrpSerAspGlySerSerLeuSerLysGluAsnTrpIleGluGluSerLys 100
 Db 274 ACNARTGGWSNGAYGNGWSNGTNAAYTAYGARAAYTNTNATHAARCAAYGCNACNAAR 333
 Qy 101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
 Db 334 AARTGYTYTGGNYTNAARAARGARACNGNTTYMGNACNTGGMNAAVGTNCAYTYACN 393
 Qy 121 GlnGlnAspProPheValCysGlu 128
 Db 394 CARCARAAYTNTTYATGTGYAAR 417
 RESULT 9
 AAD32060
 ID AAD32060 standard; DNA; 474 BP.
 XX
 AC AAD32060;
 XX
 DT 18-JUN-2002 (first entry)
 XX
 DE Pigmy rattlesnake venom gland protein Zank5, degenerate nucleic acid.
 XX Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;
 KW platelet aggregation; ds.
 XX
 XX Sistrurus miliarius.
 XX WO200214364-A2.
 XX 21-FEB-2002.
 XX 13-AUG-2001; 2001WO-US25310.
 XX 14-AUG-2000; 2000US-225072P.
 PR 14-AUG-2000; 2000US-225087P.
 PR 15-AUG-2000; 2000US-225489P.
 PR 15-AUG-2000; 2000US-225490P.
 PR 20-DEC-2000; 2000US-356997P.
 XX (ZYMO) ZYMOGENETICS INC.
 XX Sheppard PO, Bishop PD;
 XX WPI; 2002-269180/31.
 XX New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which
 PT affect blood coagulation and platelet aggregation system, useful in
 PT therapy and diagnostics, or as tools in the study of genetics or
 PT molecular biology -
 XX Disclosure; Page 78; 79pp; English.
 CC The invention relates to new pigmy rattlesnake (Sistrurus miliarius)
 CC venom gland proteins, which affect blood coagulation and platelet
 CC aggregation system. The polypeptides, which affect blood coagulation and
 CC platelet aggregation system, are useful in therapy and diagnostics. The
 CC polypeptides are also useful as an educational tool in laboratory
 CC practical kits for courses related to genetics and molecular biology,
 CC protein chemistry and antibody production and analysis. The
 CC polynucleotide or polypeptide can be used as standards or as unknowns
 CC for testing purposes. The polypeptides are also useful in identifying
 CC proteins by western blotting, protein purification, determining the
 CC weight of expressed polypeptides as a ratio to total protein expressed,
 CC identifying peptide cleavage sites, coupling amino and carboxyl terminal
 CC tags, mass spectrometry, circular dichroism to determine conformation or
 CC affinity chromatography columns to purify the protein, cloning or
 CC sequencing. The present sequence is Sistrurus miliarius venom gland
 CC protein Zank5, degenerate nucleic acid.
 XX
 SQ Sequence 474 BP; 89 A; 31 C; 84 G; 74 T; 196 other;
 Alignment Scores:
 Pred. No.: 3.41e-30 Length: 474
 Score: 329.00 Matches: 55
 Percent Similarity: 58.59% Conservative: 20
 Best Local Similarity: 42.97% Mismatches: 53
 Query Match: 46.01% Indels: 0
 DB: 24 Gaps: 0
 US-09-938-114-2 (1-129) x AAD32060 (1-474)
 Qy 1 AspCysSerSerAspTrpSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
 Db 76 AAYTGCCNWSNGNTGGTGYGNCATYAYGAYCARTAYTGYTMGMNTNATHAARMGNTN 135
 Qy 21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
 Db 136 AARACNTGGGAYGAYGCGNGARGMNTTYTGYMSNGARGCARGCNAARGGNGNCAYTYNTGN 195
 Qy 41 SerLeuGluSerSerGlyGluAlaAspPheValGlyGlnLeuLeuAlaGlnLysIleLys 60
 Db 196 WSGNTNGARAAYGAYGARGCNGTNTTYTNGCNCARYTNGTNGCNGCNAAYATHAAR 255
 Qy 61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
 Db 256 CARAAYCARTAYTAYGINTGATHGNTNMGNTNATHCARGAAYARGCNCARTGYWSN 315

QY 81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluSerLys 100
 DB 316 ACNARTCGNSGAYGGSNSNGTNSNTAYGARAAYTYNGTWAARWSNCAYWSNAAR 375
 QY 101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
 DB 376 AARTGYTYGGNYTWAARAARGACACGARTTYTNCARTGTAAYACACGAYTYGAR 435
 QY 121 GlnGlnAspPropheValCysGlu 128
 DB 436 GARAARAAYTYTNGTNGTYAAR 459
 RESULT 10
 AAI71877
 ID AAI71877 standard; DNA; 633 BP.
 AC AAI71877;
 XX
 DT 10-JAN-2002 (first entry)
 XX
 DE Snake venom blood anticoagulant halyxin B chain coding sequence.
 XX
 KW Snake; pit viper; venom; halyxin; anticoagulant; thrombogenesis;
 KW thrombosis; ds.
 XX
 OS Agkistrodon halys.
 XX
 PN KR2001049671-A.
 XX
 PD 15-JUN-2001.
 XX
 PF 29-JUN-2000; 2000KR-0036591.
 XX
 PR 29-JUN-1999; 99KR-0025105.
 XX
 PA (BIOB-) BIOBUD CO LTD.
 XX
 PI Jang YS, Jung GH, Kim DS, Koo BH, Son YD;
 XX WPI; 2001-637330/73.
 DR P-PSDB; AAM51544.
 XX
 PT Halyxin as blood anticoagulation protein separated from snake venom -
 XX Claim 1; Page 11; 21pp; Korean.
 CC The invention relates to halyxin, a novel protein with very strong
 CC blood anticoagulation activity. The protein was separated from snake
 CC venom of Agkistrodon halys brevicaudus (a Korean pit viper) and can
 CC be used in the treatment of thrombogenesis. The present sequence
 CC encodes the B chain of halyxin.
 XX
 SQ Sequence 633 BP; 179 A; 142 C; 157 G; 155 T; 0 other;
 Alignment Scores:
 Pred. No.: 2,41e-25 Length: 633
 Score: 290.00 Matches: 59
 Percent Similarity: 57.81% Conservative: 15
 Best Local Similarity: 46.09% Mismatches: 46
 Query Match: 40.56% Indels: 8
 DB: 22 Gaps: 4
 US-09-938-114-2 (1-129) x AAI71877 (1-633)
 QY 1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
 DB 93 GATTGTCCCTCTGTTGGTCTCTCTATGAAGGCAATTGCTACAAACCAITTTAATGAACA 152
 QY 21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
 DB 153 AAAAATGGGCCCGCCAGAGAAATTTCTGCACACACACACACAGCGGACATCTGGTC 212
 QY 41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60

DB 213 TCCTTCCACAGCACTGAAGACGAGATTTTGTGGTCAAGCTGGCCCTTCCAAATTTTGGC 272
 QY 61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
 DB 273 CACGGTATTTTC-----TGGATGGGACTG-----AGCAATGTCTGGAATCAATGCACG 320
 QY 81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluSerLys 100
 DB 321 TGGCAATGGAGCGAGTGTGCCAAGCTCAAAATACGAAGCGCTGGCTGGAAGAA-----TCT 374
 QY 101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
 DB 375 TATTGT-----GTCTATTTCAGTCAACTAATAACAATGGAGGAGTAGAGCTGCAG 428
 QY 121 GlnGlnAspPropheValCysGlu 128
 DB 429 ATGGAGGCATATATTGCTTGGCAG 452
 RESULT 11
 AAQ89309
 ID AAQ89309 standard; cDNA; 690 BP.
 XX
 AC AAQ89309;
 XX
 DT 25-MAR-2003 (updated)
 DT 28-NOV-1995 (first entry)
 XX
 DE Snake venom antithrombotic oligopeptide cDNA.
 XX
 KW Antithrombotic peptide; snake venom; platelet binding inhibition;
 KW von Willebrand factors; Crotalus horridus horridus; ds.
 XX
 OS Crotalus horridus horridus.
 XX
 Key Location/Qualifiers
 FH 66..515
 CDS /*tag= a
 FT
 FT
 XX WO9508573-A1.
 XX 30-MAR-1995.
 XX 21-SEP-1994; 94WO-JP01555.
 XX 22-SEP-1993; 93JP-0236975.
 XX (AJIN) AJINOMOTO KK.
 XX Fukuchi N, Iehli K, Kito M, Kobayashi T, Nagano M;
 PI Tanaka A, Yamamoto H, Yoshimoto R;
 XX WPI; 1995-139559/18.
 DR P-PSDB; AAR71981.
 XX
 PT Single-chain antithrombotic peptide - obtained by cleaving an
 PT oligopeptide from snake venom to break inter-chain di:sulphide
 PT bonds but preserve intra-chain di:sulphide bonds
 XX
 PS Example 2; Pages 47-48; 84pp; Japanese.
 XX
 CC AAR71978 and AAR71979 are snake venom derived antithrombotic peptides,
 CC specifically from the snake venom oligopeptide AAR71981, encoded by
 CC AAQ89309. These peptides have the advantage of avoiding significant
 CC thrombocytopenia when administered at the minimum dose, for in vivo
 CC inhibition of platelet von Willebrand factor binding.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 690 BP; 167 A; 180 C; 170 G; 173 T; 0 other;
 Alignment Scores:
 Pred. No.: 7.13e-25 Length: 690
 Score: 286.50 Matches: 55

Percent Similarity: 60.16% Conservative: 22
 Best Local Similarity: 42.97% Mismatches: 42
 Query Match: 40.07% Indels: 9
 DB: 16 Gaps: 2

US-09-938-114-2 (1-129) x AAQ89309 (1-690)

QY 1 AspCysSerSerAspTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
 DB 141 GAATGTCCCTCCGGTGGTTCCTATGATCGGTATGCTACAGCCCTTCAACAAGAG 200
 QY 21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
 DB 201 ATGACCTGGCCGATGTCAGAGAGTTCCTCGGAGCAGCGAAGCGGGGCATCTCTCTC 260
 QY 41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
 DB 261 TCTGTGGAACCGCCCTAGAGCATCTTGTGGACAATGTCTCTATCGCAACAAGAG 320
 QY 61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
 DB 321 TACCTCACAGCTTATCTGGATTGGACTGGAGGGTTCAAAACAAGAGCAGCCATGC--- 377
 QY 81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluSerLys 100
 DB 378 -----TCCAGCATCAGTTATGAGAACCTGGTT-----GACCCATT 413
 QY 101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
 DB 414 GAATGTTTTTATGCTGAGCAGACACAAAGCTTCGTGAGTGGTTTAAAGTTGACTGTGAA 473
 QY 121 GlnGlnAspPropheValCysGlu 128
 DB 474 CAACAACATCTTTCATATGCAAG 497

RESULT 12
 AAC61144

ID AAC61144 standard; DNA; 690 BP.

AC AAC61144;

DT 07-FEB-2001 (first entry)

DE DNA encoding a snake venom derived protein.

KW Subunit peptide production; snake venom; rattlesnake; thrombolytic;
 von Willebrand's factor; blood platelet-inhibitory activity; ds.

OS Crotalus horridus horridus.

PN WO200059926-A1.

PD 12-OCT-2000.

PF 31-MAR-2000; 2000WO-JP02127.

PR 02-APR-1999; 99JP-0096073.

PA (AJIN) AJINOMOTO CO INC.

PI Fukuchi N, Kageyama S, Kito M, Kayahara T, Yamamoto H;

DR WPI; 2000-664985/64.

DR P-FSDB; AAY85628.

PT Producing physiologically-active subunit peptides originating in
 polymer proteins by denaturation and specific separation, with lower
 antigenicity but improved solubility and stability, e.g. blood
 platelet-binding inhibitors

PS Disclosure; Page 46; 51pp; Japanese.

CC This invention relates to a method for the production of a subunit

CC peptide originating from a polymer protein with disulphide bonds within
 CC and between subunits. The method comprises denaturing the protein or its
 CC subunit using a protein denaturing agent in a solution, removing the
 CC agent in the presence of a polyoxyalkyl polyether which reacts with a
 CC thiol group and unwinds the subunit, and separating the polyoxyalkyl
 CC polyether-bound subunit peptide. The method can be used for producing
 CC physiologically-active subunit peptides for polymer proteins e.g. snake
 CC venom-originated dimer peptide with blood platelet-inhibitory activity
 CC on von Willebrand's factor. The peptides produced have platelet-binding
 CC inhibitory, and thrombolytic activity. The present sequence represents
 CC DNA encoding rattlesnake protein used in an example illustrating the
 CC method of the invention.
 XX

SQ Sequence 690 BP; 167 A; 180 C; 170 G; 173 T; 0 other;

Alignment Scores:

Pred. No.: 7, 13e-25 Length: 690
 Score: 286.50 Matches: 55
 Percent Similarity: 60.16% Conservative: 22
 Best Local Similarity: 42.97% Mismatches: 42
 Query Match: 40.07% Indels: 9
 DB: 21 Gaps: 2

US-09-938-114-2 (1-129) x AAC61144 (1-690)

QY 1 AspCysSerSerAspTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
 DB 141 GAATGTCCCTCCGGTGGTTCCTATGATCGGTATGCTACAGCCCTTCAACAAGAG 200
 QY 21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
 DB 201 ATGACCTGGCCGATGTCAGAGAGTTCCTCGGAGCAGCGAAGCGGGGCATCTCTCTC 260
 QY 41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
 DB 261 TCTGTGGAACCGCCCTAGAGCATCTTGTGGACAATGTCTATGCGCAACAAGAG 320
 QY 61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
 DB 321 TACCTCACAGCTTATCTGGATTGGACTGGAGGGTTCAAAACAAGAGCAGCCATGC--- 377
 QY 81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluSerLys 100
 DB 378 -----TCCAGCATCAGTTATGAGAACCTGGTT-----GACCCATT 413
 QY 101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
 DB 414 GAATGTTTTTATGCTGAGCAGACACAAAGCTTCGTGAGTGGTTTAAAGTTGACTGTGAA 473
 QY 121 GlnGlnAspPropheValCysGlu 128
 DB 474 CAACAACATCTTTCATATGCAAG 497

RESULT 13

AAK99834

ID AAK99834 standard; DNA; 690 BP.

AC AAK99834;

DT 19-JUL-2002 (first entry)

DE DNA encoding the antithrombotic wild-type rattlesnake protein.

KW Thrombolytic; glycoprotein Ib-binding protein; snake venom; rattlesnake;
 antithrombotic; glycoprotein Ib; long half life; low antigenicity;
 drug efficacy; gene; ds.

OS Crotalus horridus horridus.

FH Key Location/Qualifiers

FT CDS 66..515

FT /*tag= a

FT /product= "Antithrombotic wild-type rattlesnake protein"

Dd 76 GATTGTCCCTCGTGGTCTCTTATGAAGGCAATGCTACAGGCTTTCATCAACAC 135
Qy 21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
Dd 136 TCTCAGTGGCCCATCAGAGAAATTTCTGCACACTACAGCACACAGCGGACATCTGGTC 195
Qy 41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
Dd 196 TCCTCCACAGTACTCAAGACAGAGATTTTGTCAGCTGGCTCCCTCCAAATTTTGGC 255
Qy 61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
Dd 256 CACGGTATTTTC-----TGGATGGGACTG-----AGCAATGCTCGGAATCAATGCAC 303
Qy 81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluSerLys 100
Dd 304 TGGCAATGAGCAATGCTCCAGCTCAATAGTCAAGCTGGGCTGAG----- 351
Qy 101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
Dd 352 ---ATCTATTGTGCTATTTCAGTCAACTAATAACAATAGGAGGATAGACCTGCAGA 408
Qy 121 GlnGlnAspPropheValCysGlu 128
Dd 409 ATGGAGGCATATTTCGCTGCGAG 432

RESULT 15

AAD32053
ID AAD32053 standard; DNA; 483 BP.
XX
AC AAD32053;
XX

18-JUN-2002 (first entry)

Pigmy rattlesnake venom gland protein, Zsnk2 gene.

Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;
platelet aggregation; gene; Zsnk2; ds.

Sistrurus miliarius.

Key Location/Qualifiers
CDS 3..458
FT /*tag= a
FT /product= "Zsnk2 protein"
FT /note= "CDS does not include start codon"
FT /partial
FT sig_peptide 3..59
FT /*tag= b
FT mat_peptide 60..455
FT /*tag= c
FT /product= "Mature Zsnk2 protein"

WO200214364-A2.

21-FEB-2002.

13-AUG-2001; 2001WO-US25310.

14-AUG-2000; 2000US-225072P.

14-AUG-2000; 2000US-225087P.

15-AUG-2000; 2000US-225489P.

15-AUG-2000; 2000US-225490P.

20-DEC-2000; 2000US-356997P.

(Zymo) ZYMOGENETICS INC.

Sheppard PO, Bishop PD;

WPI; 2002-269180/31.

P-PSDB; AAE20176.

New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which

PT affect blood coagulation and platelet aggregation system, useful in
PT therapy and diagnostics, or as tools in the study of genetics or
PT molecular biology
XX
PS Claim 5; Page 71-72; 79pp; English.

CC The invention relates to new pigmy rattlesnake (Sistrurus miliarius)
CC venom gland proteins, which affect blood coagulation and platelet
CC aggregation system. The polypeptides, which affect blood coagulation and
CC platelet aggregation system, are useful in therapy and diagnostics. The
CC polypeptides are also useful as an educational tool in laboratory
CC practical kits for courses related to genetics and molecular biology,
CC protein chemistry and antibody production and analysis. The
CC polynucleotide or polypeptide can be used as standards or as unknowns
CC for testing purposes. The polypeptides are also useful in identifying
CC proteins by western blotting, protein purification, determining the
CC weight of expressed polypeptides as a ratio to total protein expressed,
CC identifying peptide cleavage sites, coupling amino and carboxyl terminal
CC tags, mass spectrometry, circular dichroism to determine conformation or
CC affinity chromatography columns to purify the protein, cloning or
CC sequencing. The present sequence is Sistrurus miliarius venom gland
CC protein, Zsnk2 gene.

SQ Sequence 483 BP; 129 A; 112 C; 123 G; 119 T; 0 other;

Alignment Scores:

Pred. No.: 2.21e-22 Length: 483
Score: 264.00 Matches: 55
Percent Similarity: 57.03% Conservative: 18
Best Local Similarity: 42.97% Mismatches: 49
Query Match: 36.92% Indels: 6
DB: 24 Gaps: 3

US-09-938-114-2 (1-129) x AAD32053 (1-483)

Qy 1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
Dd 60 GATTGTCCCTCTGACTGGTCTCTCTATGATGAGCATTCCTACAGGCTTTCAGTGAATC 119
Qy 21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
Dd 120 AAACCTGGGATGATGACAGAGATTTCTGCTACACACAGACAGAGCGCCCTGGCC 179
Qy 41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
Dd 180 TCCATCCACAGCAGTGAAGAGAAAGCTTTTGTGGCAAACTGGCTCCCAACTTGAA 239
Qy 61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
Dd 240 TTCACCTCC-----ATGTGGATCGGACTGAAA-----GATCTATGGAAGAATGCAAA 287
Qy 81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluSerLys 100
Dd 288 TGGCAGTGGAGCGATGACACCAAACTGGACTACAAAGCCTGGACTCGAGA-----CCC 341
Qy 101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
Dd 342 TATTGTACAGTAATAGTAGTCAAGACAGATAGGATCTTTTGGTTCAATAGAGTTGCGAA 401
Qy 121 GlnGlnAspPropheValCysGlu 128
Dd 402 AAGACTATCTTTTGTCTGCAAG 425

Search completed: December 8, 2003, 14:04:46
Job time : 446.023 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 8, 2003, 13:49:04 ; Search time 101.726 Seconds
(without alignments)
559.725 Million cell updates/sec

Title: US-09-938-114-2

Perfect score: 715

Sequence: 1 DCSWSSVEGHCKYFKQS.....GPHKWNFYCEQDPFVCEA 129

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlh

-Q=/cgn2_1/USPIO_spool/US09938114/runat_08122003_091001_22878/app_query.fasta_1.725

-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPEL=0

-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi

-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15

-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US09938114@cgn_1_1_95@runat_08122003_091001_22878 -NCPU=6 -ICPU=3

-NO_WMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG

-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA.*

1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*

2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*

3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*

4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*

5: /cgn2_6/ptodata/1/ina/PTUS COMB.seq.*

6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	653	91.3	544	4	US-09-058-740-1
2	286.5	40.1	690	2	US-08-612-840A-7
3	216	30.2	454	4	US-09-058-740-12
4	201.5	28.2	777	3	US-09-146-969-1
5	201.5	28.2	4771	3	US-08-840-062-3
6	197.5	27.6	4588	3	US-08-454-062-1
7	195.5	27.3	590	2	US-08-454-557C-32
8	195.5	27.3	590	2	US-08-340-426D-32
9	195.5	27.3	590	2	US-08-450-673C-32
10	195.5	27.3	590	5	PCT-US95-17111A-32
11	184.5	25.8	5191	1	US-08-340-428B-1
12	184.5	25.8	5191	5	PCT-US93-07306-1

13	183	25.6	522	1	US-07-778-156-4	Sequence 4, Appli
14	183	25.6	522	2	US-08-422-166-4	Sequence 4, Appli
15	183	25.6	797	2	US-08-464-637-1	Sequence 1, Appli
16	183	25.6	797	2	US-08-822-261-5	Sequence 5, Appli
17	183	25.6	797	4	US-09-226-852-5	Sequence 5, Appli
18	183	25.6	798	1	US-07-778-156-12	Sequence 12, Appli
19	183	25.6	798	2	US-08-822-261-6	Sequence 6, Appli
20	183	25.6	798	3	US-08-422-166-12	Sequence 12, Appli
21	183	25.6	798	3	US-09-146-969-2	Sequence 2, Appli
22	183	25.6	798	4	US-09-226-852-6	Sequence 6, Appli
23	181.5	25.4	272	2	US-08-612-840A-6	Sequence 6, Appli
24	181.5	25.4	614	2	US-08-729-103-2	Sequence 2, Appli
25	181.5	25.4	1114	2	US-08-468-413-1	Sequence 1, Appli
26	181.5	25.4	1114	3	US-09-162-508-1	Sequence 1, Appli
27	181.5	25.4	1114	5	PCT-US95-07169-1	Sequence 4, Appli
28	175	24.5	558	1	US-08-909-725-4	Sequence 3, Appli
29	175	24.5	586	3	US-09-146-969-3	Sequence 3, Appli
30	175	24.5	747	2	US-08-401-530A-1	Sequence 1, Appli
31	175	24.5	747	2	US-08-709-662-1	Sequence 2, Appli
32	175	24.5	762	2	US-08-822-261-2	Sequence 2, Appli
33	175	24.5	762	4	US-09-226-852-2	Sequence 2, Appli
34	168	23.5	4978	1	US-08-220-603A-1	Sequence 1, Appli
35	163	22.8	885	1	US-08-365-103B-3	Sequence 3, Appli
36	163	22.8	924	1	US-08-365-103B-5	Sequence 5, Appli
37	163	22.8	1005	1	US-08-365-103B-1	Sequence 1, Appli
38	161.5	22.6	8224	6	5180808-1	Patent No. 5180808
39	159	22.2	474	1	US-07-778-156-8	Sequence 8, Appli
40	159	22.2	474	2	US-08-422-166-8	Sequence 8, Appli
41	159	22.2	793	1	US-07-778-156-1	Sequence 1, Appli
42	159	22.2	793	2	US-08-422-166-1	Sequence 1, Appli
43	157.5	22.0	3259	5	PCT-US95-03747-1	Sequence 2, Appli
44	155	21.7	5169	4	US-09-194-612A-2	Sequence 2, Appli
45	144	20.1	2318	4	US-09-620-312D-733	Sequence 733, App

ALIGNMENTS

RESULT 1

US-09-058-740-1

; Sequence 1, Application US/09058740

; Patent No. 6489451

; GENERAL INFORMATION:

; APPLICANT: Chun Wang, Benjamin XY Li, Xin Cheng,

; Jing Liu, Li-Wen Niu, Wan-Zhi Huang,

; Zhen-Yu Xu, Dan Luo, Lian-Di Kang,

; Jin-Guo Ding, Fang Rong, Yan Liu and

; Hui-Ran Chen

; TITLE OF INVENTION: AN ANTI-THROMBOSIS ENZYME FROM THE SNAKE

; VENOM OF AGKISTRODON ACUTUS

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Lyon & Lyon

; STREET: 633 West Fifth Street

; CITY: Suite 4700

; STATE: Los Angeles

; COUNTRY: U.S.A.

; ZIP: 90071-2066

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

; storage

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: IBM P.C. DOS 5.0

; SOFTWARE: Fastseq for Windows 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/058,740

; FILING DATE: 10-Apr-1998

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: <Unknown>

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Chen, Anthony C.

US-09-058-740-12
; Sequence 12, Application US/09058740
; Patent No. 6489451
; GENERAL INFORMATION:
; APPLICANT: Chun Wang, Benjamin XY Li, Xin Cheng,
; Jing Liu, Li-Wen Niu, Wan-Zhi Huang,
; Zhen-Yu Xu, Dan Luo, Lian-Di Kang,
; Jin-Guo Ding, Fang Rong, Yan Liu and
; Hui-Ran Chen
; TITLE OF INVENTION: AN ANTITHROMBOSIS ENZYME FROM THE SNAKE
; VENOM OF AGKISTRODON ACUTUS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA: US/09/058,740
; FILING DATE: 10-Apr-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Chen, Anthony C.
; REGISTRATION NUMBER: 38,673
; REFERENCE/DOCKET NUMBER: 233/298
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 454 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 3...440
; OTHER INFORMATION: "N" stands for any base.
; "Xaa" stands for any amino acid.
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:

Alignment Scores:
Pred. No.: 6,876-19 Length: 454
Score: 216.00 Matches: 50
Percent Similarity: 49.63% Conservative: 17
Best Local Similarity: 37.04% Mismatches: 46
Query Match: 30.21% Indels: 22
DB: 4 Gaps: 6

US-09-938-114-2 (1-129) x US-09-058-740-12 (1-454)

Qy 1 AspCysSerAspTrpSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
Db 72 GATTGTCCCTCTGAGTGTCTCTCTATGAGGGCAATTCTACAGCCCTTCATGAACCT 131
Qy 21 LysThrTrpAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeu--- 39
Db 132 AAGACCTGGGAGATGACAGAAATTTCTGCACACAAACAAAGGAGCCATCTGCT 191

Qy 40 -----ValSerIleGluSerSerGlyGluAlaAspPheValGlyGln 53
Db 192 CTCACAGCAGTCAGAGCGATTCTGTNNNNNNNNNTGGTCACGTTGACCACACC----- 245
Qy 54 LeuIleAlaGlnLysIleLysSerAlaLysIleHisValTrpIleGlyLeuArgAlaGln 73
Db 246 -----AAGTTGAACTGATTAGTCTG-----ATTGGACTGAAG----- 278
Qy 74 AsnLysGluLysGlnCysSerIleGluTrpSerAspGlySerSerIleSerLysGluAsn 93
Db 279 AACATCTGAACCGATGCTACTGGAAGTGGACGATGGCACCAGCTCAGCTACAAAGAC 338
Qy 94 TrpIleGluGluGluSerLysLysCysLeuGlyValHisIleGluThrGlyPheHisLys 113
Db 339 TGG-----CGTGAACAATTTGATGTCTC-----GTATCCAGGACAGTAAATAACGAA 386
Qy 114 TrpGluAsnPheTyrCysGluGlnGlnAspProPheValCysGlu 128
Db 387 TGGCTAAGTATGGACTGCGGCACCTACTTGTCTTTCGTCTGCAAG 431

RESULT 4

US-09-146-969-1
; Sequence 1, Application US/09146969
; Patent No. 6228585
; GENERAL INFORMATION:
; APPLICANT: Dieckgrafe, Brian K.
; TITLE OF INVENTION: Gene Markers for Chronic Mucosal Injury
; FILE REFERENCE: 04255.75314
; CURRENT APPLICATION NUMBER: US/09/146,969
; CURRENT FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-146-969-1

Alignment Scores:
Pred. No.: 1,07e-16 Length: 777
Score: 201.50 Matches: 44
Percent Similarity: 53.03% Conservative: 26
Best Local Similarity: 33.33% Mismatches: 53
Query Match: 28.18% Indels: 9
DB: 3 Gaps: 5

US-09-938-114-2 (1-129) x US-09-146-969-1 (1-777)

Qy 2 CysSerSerAspTrpSerTyrGluGlyHisCysTyrLysValPheLysGlnSerLys 21
Db 188 TGCCAGAAAGGCACCAATGCTATCGTCTCTACTCTACTTAAATGACACCGTGAG 247
Qy 22 ThrTrpThrAspAlaGluSerPheCysThrLysGlnValAenGlyGlyHisLeuValSer 41
Db 248 ACCTGGTTGATGAGATCTCTATTGC---CAGAACATGAATCGGCAACCTGGTGCT 304
Qy 42 IleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLysSer 61
Db 305 GTGCTCACCAGCGGAGGGTGCCTTTGTGCCTCCTCATGATTAGGAGAGT---GGCACT 361
Qy 62 AlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSerIle 81
Db 362 GATGACTTCAATGTCTGGATTGGCTCCATGACCCCAAAAGAACCCGCGCTGG----- 415
Qy 82 GluTrpSerAspGlySerSerIleSerLysGluAsnTrp---IleGluGluGluSerLys 100
Db 416 CACTGGAGCAGTGGGTCCCTGCTCTCTACAAAGTCCTGGGCGCATTTAGGCCCAAGCAGT 475
Qy 101 -----LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsn 116
Db 476 GTTAACTCTGGCTACTGTGTGAGCTGACCTCAAGCAGAGATTCCAGAAATGGAAGAT 535
Qy 117 PheTyrCysGluGlnGlnAspProPheValCysGlu 128


```

Db 1416 -----GAGCTAGGATGGCTC-----AATGATTGAACCTGCAGATGAAT 1457
QY 81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrp----- 94
Db 1458 TTTGAGTGCTCCGAGGGAGCCTCGTAGCTTCAACCCACTGGCACCCTTTGAGCCCAAC 1517
QY 95 ---IleGluGluGluSerLysLysCysLeuGlyValHisIleGluThrGlyPheHisLys 113
Db 1518 AACTTTGTCGACGCTGGAGGACTGTGTACATCTGGGGCCCGGAAGGA-----CGC 1571
QY 114 TrpGluAsnPheTyrcysgluInGlnAspPropheValCysGlu 128
Db 1572 TGGACGACAGCAGTCCCTGTGAACCAAGTCCITGGCCATCCATTGCAAG 1616

RESULT 7
US-08-454-557C-32
; Sequence 32, Application US/08454557C
; Patent No. 5830670
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,557C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 590 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
US-08-454-557C-32

Alignment Scores:
Pred. No.: 4.24e-16 Length: 590
Score: 195.50 Matches: 43
Percent Similarity: 53.97% Conservative: 25
Best Local Similarity: 34.13% Mismatches: 49
Query Match: 27.34% Indels: 9
DB: 2 Gaps: 5

US-09-938-114-2 (1-129) x US-08-454-557C-32 (1-590)
QY 8 SerTyrGluGlyHisCysTyrLysValPheLysGlnSerLysThrTrpThrAspAlaGlu 27
Db 3 GCCTATCGCTCCTACTACTACTTAAATGAACCGTAGACCTGGTGTGATGCAGAT 62
QY 28 SerPheCysThrLysGlnValAsnGlyGlyHisLeuValSerIleGluSerSerGlyGlu 47
Db 63 CTCTATTGTC---CAGAATCAATTCGGGCAACCTGGTGTCTGTGCTCACCCAGCCGCGAG 119

```

```

QY 48 AlaAspPheValGlyGlnLeuIleAlaGlnLysIleLysSerAlaLysIleHisValTyr 67
Db 120 GGTGCTTTGTGGCTCACTGATTAAAGGAGACT---GGCACTGATCACTTCAATGTCTGG 176
QY 68 IleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSerIleGluTrpSerAspGlySer 87
Db 177 APTGCTCTCCATGACCCCAAAAGAACCGCCCTGG-----CACTGGAGCAGTGGGTCC 230
QY 88 SerIleSerLysGluAsnTrp---IleGluGluGluSerLys-----LysCys 102
Db 231 CTGGTCTCTCAAGTCTCTGGGCATTGGAGCCCAAGCAGTGTAAATCCTGGCTACTGT 290
QY 103 LeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrcysgluInGln 122
Db 291 GTGAGCCTGACCTCAAGCACAGGATTCAGAAATGAAGGATGTGCTTGTGAACACAAAG 350
QY 123 AspPropheValCysGlu 128
Db 351 TTCTCCTTTGTGTGCAAG 368

RESULT 8
US-08-340-426D-32
; Sequence 32, Application US/08340426D
; Patent No. 5948634
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340,426D
; FILING DATE: 14-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 590 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
US-08-340-426D-32

Alignment Scores:
Pred. No.: 4.24e-16 Length: 590
Score: 195.50 Matches: 43
Percent Similarity: 53.97% Conservative: 25
Best Local Similarity: 34.13% Mismatches: 49
Query Match: 27.34% Indels: 9
DB: 2 Gaps: 5

US-09-938-114-2 (1-129) x US-08-340-426D-32 (1-590)
QY 8 SerTyrGluGlyHisCysTyrLysValPheLysGlnSerLysThrTrpThrAspAlaGlu 27

```

```

Db 3 GCCTATCGCTCTACTGCTACTTAAATGAAGACCGCTGAGACCTGGTGTGATGCAGAT 62
Qy 28 SerPheCysThrLysGlnValAsnGlyHisLeuValSerIleGluSerSerGlyGlu 47
Db 63 CTCATATGCG---CAGACATGAATTCGGGCAACCTGGTGTCTGTGCTCACCAGCCGCGAG 119
Qy 48 AlaAspPheValGlyGlnLeuIleAlaGlnLysIleLysSerAlaLysIleHisValTrp 67
Db 120 GGTGCTTTGTGGCTCTACTGATTAAGGAGAGT---GGCACTGATGAGCTTCAATGCTGTGG 176
Qy 68 IleGlyLeuArgAlaGlnAsnLysGlnCysSerIleGluTrpSerAspGlySer 87
Db 177 ATTGGCTTCCATGACCCCAAAAGAACCGCGCTGG-----CACTGGAGCAGTGGGTCC 230
Qy 88 SerIleSerLysGluAsnTrp---IleGluGluGluSerLys-----LysCys 102
Db 231 CTGGTCTCTACAGTCTCTGGGCGATTGGAGCCCAAGCAGTGTAAATCCTGGCTACTGT 290
Qy 103 LeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrcysGluGlnGln 122
Db 291 GTGAGCCTGACCTCAAGCACAGGATTCCAGAAATGGAAGGATGTGCCTTGTGAAGACAAG 350
Qy 123 AspPropheValCysGlu 128
Db 351 TTCCTCTTTGTCTGCAAG 368

```

RESULT 9

```

US-08-450-673C-32
; Sequence 32, Application US/08450673C
; Patent No. 5948888
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wanda, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: Of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,673C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 590 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
US-08-450-673C-32
Alignment Scores:
Pred. No.: 4.24e-16 Length: 590
Score: 195.50 Matches: 43
Percent Similarity: 53.97% Conservative: 25
Best Local Similarity: 34.13% Mismatches: 49

```

```

Query Match: 27.34% Indels: 9
DB: Gaps: 5
US-09-938-114-2 (1-129) x US-08-450-673C-32 (1-590)
Qy 8 SerTyrgluGlyHisCysValPheLysGlnSerLysThrTrpThrAspAlaGlu 27
Db 3 GCCTATCGCTCTACTGCTACTTAAATGAAGACCGCTGAGACCTGGTGTGATGCAGAT 62
Qy 28 SerPheCysThrLysGlnValAsnGlyHisLeuValSerIleGluSerSerGlyGlu 47
Db 63 CTCATATGCG---CAGACATGAATTCGGGCAACCTGGTGTCTGTGCTCACCAGCCGCGAG 119
Qy 48 AlaAspPheValGlyGlnLeuIleAlaGlnLysIleLysSerAlaLysIleHisValTrp 67
Db 120 GGTGCTTTGTGGCTCTACTGATTAAGGAGAGT---GGCACTGATGAGCTTCAATGCTGTGG 176
Qy 68 IleGlyLeuArgAlaGlnAsnLysGlnCysSerIleGluTrpSerAspGlySer 87
Db 177 ATTGGCTTCCATGACCCCAAAAGAACCGCGCTGG-----CACTGGAGCAGTGGGTCC 230
Qy 88 SerIleSerLysGluAsnTrp---IleGluGluGluSerLys-----LysCys 102
Db 231 CTGGTCTCTACAGTCTCTGGGCGATTGGAGCCCAAGCAGTGTAAATCCTGGCTACTGT 290
Qy 103 LeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrcysGluGlnGln 122
Db 291 GTGAGCCTGACCTCAAGCACAGGATTCCAGAAATGGAAGGATGTGCCTTGTGAAGACAAG 350
Qy 123 AspPropheValCysGlu 128
Db 351 TTCCTCTTTGTCTGCAAG 368

```

RESULT 10

```

PCT-US95-17111A-32
; Sequence 32, Application PC/TUS9517111A
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wanda, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and
; TITLE OF INVENTION: Detection of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/17111A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/340,426
; FILING DATE: 14-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 590 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both

```

```

; TOPOLOGY: both
PCT-US95-17111A-32

Alignment Scores:
Pred. No.: 4,24e-16 Length: 590
Score: 195.50 Matches: 43
Percent Similarity: 53.97% Conservative: 25
Best Local Similarity: 34.13% Mismatches: 49
Query Match: 27.34% Indels: 9
DB: 5 Gaps: 5

US-09-938-114-2 (1-129) x PCT-US95-17111A-32 (1-590)

QY 8 SerTyrGluGlyHisCysTyrLysValPheLysGlnSerLysThrTrpThrAspAlaGlu 27
DB 3 GCCTATGCTCTACTGCTACTACTTAAATGAAGACCGCTGAGACCTGGGTTGATGCAGAT 62
QY 28 SerPheCysThrLysGlnValAsnGlyGlyHisLeuValSerLleGluSerSerGlyGlu 47
DB 63 CTCATTCG---CAGAAATGAATTCGGGCAACCTGGTGTCTGCTCACCAGGCCGAG 119
QY 48 AlaAspValGlyGlnLeuLeuAlaGlnLysLysSerAlaLysLleHisValTrp 67
DB 120 GGTGCTTTGTGGCTCTACTGATTAAGGAGCT---GGCACTGATGACTTCAATGCTCGG 176
QY 68 IleGlyLeuArgAlaGlnAsnLysGlnLysSerLleGluTrpSerAspGlySer 87
DB 177 ATTGGCCCTCCATGACCCCAAAAGAACCCGCGCTGG-----CACTGGAGCAGTGGGTCC 230
QY 88 SerLysSerLysGluAsnTrp---IleGluGluGluSerLys-----LysCys 102
DB 231 CTGGTCTCTCAAGTCTCTGGGCAATGGAGCCCAAGCAGTGTAATCCTGGCTACTGT 290
QY 103 LeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGluGln 122
DB 291 GTGAGCCTGACCTCAAGCACAGGATTCAGAAATGGAAGGATGTCCTTTGTGAACACAAG 350
QY 123 AspProPheValCysGlu 128
DB 351 TTCTCCTTTGTCTGCAAG 368

RESULT 11
US-08-340-428B-1
; Sequence 1, Application US/08340428B
; Patent No. 5648465
; GENERAL INFORMATION:
; APPLICANT: MARGOLIS, Richard U.
; APPLICANT: RAUCH, Uwe
; APPLICANT: MARGOLIS, Renee K.
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES FOR A
; TITLE OF INVENTION: NEUROCAN AS A CHONDROITIN SULFATE PROTEOGLYCAN
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340,428B
; FILING DATE: 14 No. 5648465ember 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/922,911
; FILING DATE: 03 August 1992
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:

; NAME: Browdy, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: Margolis=1A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5191 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 77..3847
US-08-340-428B-1

Alignment Scores:
Pred. No.: 2,39e-13 Length: 5191
Score: 184.50 Matches: 39
Percent Similarity: 45.93% Conservative: 23
Best Local Similarity: 28.89% Mismatches: 52
Query Match: 25.80% Indels: 21
DB: 1 Gaps: 6

US-09-938-114-2 (1-129) x US-08-340-428B-1 (1-5191)

QY 2 CysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSerLys 21
DB 3161 TGGACCGCTGGCTGGCACAATTCAGGGCCCACTGCTACCGCTACTTTGCTCATCGCGG 3220
QY 22 ThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuValSer 41
DB 3221 GCCTGGAGAGACGACAGAGAGACTGTCAGGCGCGCA-----GCCGCGCACCTGACAAGT 3274
QY 42 IleGluSerSerGlyGluAlaAspPheValGlyGlnLeuLeuAlaGlnLysLleLysSer 61
DB 3275 GTCCACTCCCAAGAGACGACCAAGTTTATTAACTGTTT-----3313
QY 62 AlaLysLleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSerLle 81
DB 3314 GGACACAGAAATTCATGATTTGGCCCTGATGACAGCAGTAGAGAG-----GACTTC 3367
QY 82 GluTrpSerAspGlySerSerLysGluAsnTrpIleGluGluGlu-----98
DB 3368 CAGTGGACAGACAACACAGGACTGCAATATGAGAACTGGAGAGAGAGCAGCGGATAAT 3427
QY 99 -----SerLysLysCys---LeuGlyValHisLleGluThrGlyPheHisLys 113
DB 3428 TTCTTCGAGGTGGGAGGATTTGTGTGTGATGGTGGCGCATGAGAAATGGA-----CGC 3481
QY 114 TrpGluAsnPheTyrCysGluGlnGlnAspProPheValCysGlu 128
DB 3482 TGGATGATGTCCTCCTGTAACCTACACCTCCCTACGTCTGCAAG 3526

RESULT 12
PCT-US93-07306-1
; Sequence 1, Application PC/TUS9307306
; GENERAL INFORMATION:
; APPLICANT: MARGOLIS, Richard U.
; APPLICANT: RAUCH, Uwe
; APPLICANT: MARGOLIS, Renee K.
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES FOR A
; TITLE OF INVENTION: NEUROCAN AS A CHONDROITIN SULFATE PROTEOGLYCAN
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20004

```



```

FEATURE:
; NAME/KEY: CDS
; LOCATION: 43..567
; OTHER INFORMATION: /product= "human"
; OTHER INFORMATION: pancreatitis-associated protein"
; OTHER INFORMATION: /note= "see, Fig. 3"
US-08-464-637-1

```

```

Alignment Scores:
Pred. No.: 2.6e-14 Length: 797
Score: 183.00 Matches: 43
Percent Similarity: 48.89% Conservative: 23
Best Local Similarity: 31.85% Mismatches: 59
Query Match: 25.59% Indels: 10
DB: 2 Gaps: 5

```

US-09-938-114-2 (1-129) x US-08-464-637-1 (1-797)

```

Qy 2 CysSerSerAspTrpSerSerTyrluGlyHisCysTyrluValPheLysGlnSerLys 21
Db 160 TGTCCCAAGGCTCCAGGCTATGGCTCCACATGCTATGCTTTTGTACCAAAA 219
Qy 22 ThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuValSer 41
Db 220 TCCTGGACAGATGCAGATCTGGCTGCCAGAGCGGCCCTCT--GGAAACCTGGTGTCT 276
Qy 42 IleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLysSer 61
Db 277 GTGCTCAGTGGGCTGAGGGATCCTTCGTCTCCTCCTGGTG--AAGAGCATTTGGTAAC 333
Qy 62 AlaLysIleHisValTrpIleGlyLeuArg-----AlaGlnAsnLysGluLysGlnCys 79
Db 334 AGTACTCATAGTCTGGATGGGCTCCATGACCCACACAGGGCCAGCCCAATGGA 393
Qy 80 ---SerIleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGlu 98
Db 394 GNAGTTGGAGTGGAGTAGCAGTGTATGATGATTAATTTTGCATGGGAGAGAAATCCC 453
Qy 99 SerLys-----LysCysLeuGlyValHisIleGluThrGlyPheHisLys 113
Db 454 TCACCATCTCAAGCCCGCCACTGTGCGAGCTGTGAGAGACACAGCATTTCTGAGG 513
Qy 114 TrpGluAsnPheTyrluGlnGlnAspProPheValCysGlu 128
Db 514 TGGAAAGATTATACTGTAATGTGAGTTACCTTAIGTCTGCAAG 558

```

Search completed: December 8, 2003, 17:27:00
Job time : 106.726 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 8, 2003, 16:08:50 ; Search time 513.051 Seconds
(without alignments)
835.675 Million cell updates/sec

Title: US-09-938-114-2

Perfect score: 715

Sequence: 1 DCSDDSSVEGHCYKVFQKS.....GFKWENFYCEQDDFVCEA 129

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2201672 seqs, 1661799599 residues

Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09938114/runat_08122003_091002_22963/app_query.fasta_1.725
-DB=Published Applications NA -QFM=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LCOEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09938114 @CGN1.1.504 @runat_08122003_091002_22963
-NCPU=6 -ICPU=3 -NO WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA.*

1:	/cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2:	/cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3:	/cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4:	/cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5:	/cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6:	/cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7:	/cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8:	/cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9:	/cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10:	/cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11:	/cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12:	/cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13:	/cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
14:	/cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
15:	/cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
16:	/cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
17:	/cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
18:	/cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description

1	653	91.3	544	11	US-09-938-114-1	Sequence 1, Appli
2	477	66.7	721	10	US-09-929-230-4	Sequence 4, Appli
3	477	66.7	721	13	US-10-226-420-4	Sequence 4, Appli
4	434	60.7	580	13	US-09-929-230-7	Sequence 7, Appli
5	434	60.7	580	13	US-10-226-420-7	Sequence 7, Appli
6	397	55.5	456	13	US-09-929-230-6	Sequence 6, Appli
7	397	55.5	456	13	US-10-226-420-6	Sequence 6, Appli
8	387	54.1	725	13	US-09-929-230-10	Sequence 10, Appl
9	387	54.1	725	13	US-10-226-420-10	Sequence 10, Appl
10	378	52.9	432	10	US-09-929-230-9	Sequence 10, Appl
11	378	52.9	432	13	US-10-226-420-9	Sequence 10, Appl
12	329	46.0	474	13	US-09-929-230-12	Sequence 9, Appli
13	329	46.0	474	13	US-10-226-420-12	Sequence 12, Appli
14	286.5	40.1	690	10	US-09-969-763-2	Sequence 12, Appl
15	264	36.9	483	13	US-09-929-230-1	Sequence 1, Appli
16	264	36.9	483	13	US-10-226-420-1	Sequence 1, Appli
17	210	29.4	453	10	US-09-929-230-3	Sequence 1, Appli
18	210	29.4	453	13	US-10-226-420-3	Sequence 3, Appli
19	201.5	28.2	562	9	US-09-764-870-174	Sequence 3, Appli
20	201.5	28.2	562	15	US-10-125-540-174	Sequence 174, App
21	201.5	28.2	777	9	US-09-739-262-1	Sequence 1, Appli
22	201.5	28.2	777	13	US-10-133-937-85	Sequence 85, Appl
23	201.5	28.2	800	13	US-09-997-003-11	Sequence 11, Appl
24	201.5	28.2	843	9	US-09-925-301-340	Sequence 340, App
25	201.5	28.2	843	13	US-09-997-003-24	Sequence 24, Appl
26	201.5	28.2	1066	9	US-09-764-870-15	Sequence 15, Appl
27	201.5	28.2	1066	15	US-10-125-540-15	Sequence 15, Appl
28	201.5	28.2	5641	15	US-10-015-219-1733	Sequence 1733, Ap
29	201.5	28.2	5709	15	US-10-015-219-1734	Sequence 1734, Ap
30	201.5	28.2	6214	15	US-10-015-219-1737	Sequence 1737, Ap
31	193.5	27.1	710	15	US-10-060-036-498	Sequence 498, App
32	192.5	26.9	539	15	US-10-060-036-3668	Sequence 3668, Ap
33	192.5	26.9	541	15	US-10-060-036-4035	Sequence 4035, Ap
34	192.5	26.9	554	15	US-10-060-036-3362	Sequence 3362, Ap
35	192.5	26.9	567	15	US-10-060-036-3129	Sequence 3129, Ap
36	192.5	26.9	637	15	US-10-060-036-3637	Sequence 3637, Ap
37	192.5	26.9	644	15	US-10-060-036-3764	Sequence 3764, Ap
38	192.5	26.9	767	10	US-09-920-300A-1748	Sequence 1748, Ap
39	192.5	26.9	767	13	US-10-099-926-1748	Sequence 1748, Ap
40	192.5	26.9	767	14	US-10-033-528-1748	Sequence 1748, Ap
41	192.5	26.9	845	9	US-09-925-237-294	Sequence 294, App
42	189.5	26.5	548	15	US-10-060-036-3709	Sequence 3709, Ap
43	188.5	26.4	534	15	US-10-060-036-3517	Sequence 3517, Ap
44	188.5	26.4	642	10	US-09-920-300A-1759	Sequence 1759, Ap
45	188.5	26.4	642	13	US-10-099-926-1759	Sequence 1759, Ap

ALIGNMENTS

RESULT 1
US-09-938-114-1
; Sequence 1, Application US/09938114
; Publication No. US20030022350A1
; GENERAL INFORMATION:
; APPLICANT: Chun Wang, Benjamin XY Li, Xin Cheng,
; Jing Liu, Li-Wen Niu, Wan-Zhi Huang,
; Zhen-Yu Xu, Dan Luo, Lian-Di Kang,
; Jin-Guo Ding, Fang Rong, Yan Liu and
; Hui-Ran Chen
; TITLE OF INVENTION: AN ANTIHROMBOSIS ENZYME FROM THE SNAKE
; VENOM OF AGKISTRODON ACUTUS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; storage

LOCATION: (0)...(0)
OTHER INFORMATION: Zsnk3
US-10-226-420-4

Alignment Scores:

Pred. No.: 2,41e-57 Length: 721
Score: 477.00 Matches: 83
Percent Similarity: 76.74% Conservative: 16
Best Local Similarity: 64.34% Mismatches: 30
Query Match: 66.71% Indels: 0
DB: 13 Gaps: 0

US-09-938-114-2 (1-129) x US-10-226-420-4 (1-721)

QY 1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
Db 160 GATTGTCCTCGGTGGTCCCTATGATGACGATGCTACAGGCTCTCAACAACTC 219
QY 21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
Db 220 AAGACGTGGGACGATGCGAGAGGTTCTGCTCGGAGCAGCGGAGCGGGCATCTCGTC 279
QY 41 SerLeuGluSerSerGlyGluAlaAspPheValGlyGlnLeuLeuAlaGlnLysLys 60
Db 280 TCTATCGAAGCTCCGAAAGCAGCAGCTTTGTGGCCAGCTGGTCCCTGAGAACAGGAGG 339
QY 61 SerAlaLysLeuHisValTrpLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
Db 340 AGAGCAATCTCTATATCTGATCGGACTGAGGGTTCAAGGCAAGAGAAAGCAATGCGAC 399
QY 81 IleGluTrpSerAspGlySerSerLeuSerLysGluAsnTrpIleGluGluSerLys 100
Db 400 CGAAGTGGAGCGATGGCTCCAGCGCTCAGTATGAGAACTGGATGAGCAATCCAAA 459
QY 101 LysCysLeuGlyValHisLeuGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
Db 460 ACATGTCTTTGGGCTGCAACAGGCAAAATATCATAGTGGTCAATATTACTGTGA 519
QY 121 GlnGlnAspPropheValCysGluAla 129
Db 520 GAATAAATCTTTTGTCTGCGAGGCA 546

RESULT 4

US-09-929-230-7
Sequence 7, Application US/09929230
Patent No. US20020161203A1
GENERAL INFORMATION:
APPLICANT: Sheppard, Paul O.
APPLICANT: Bishop, Paul D.
TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
FILE REFERENCE: 00-72
CURRENT APPLICATION NUMBER: US/09/929,230
CURRENT FILING DATE: 2001-08-13
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 580
TYPE: DNA
ORGANISM: Sistrurus miliarius
FEATURE:
NAME/KEY: CDS
LOCATION: (3)...(434)
NAME/KEY: misc feature
LOCATION: (0)...(0)
OTHER INFORMATION: Zsnk4
US-09-929-230-7

Alignment Scores:

Pred. No.: 2.04e-51 Length: 580
Score: 434.00 Matches: 75
Percent Similarity: 74.22% Conservative: 20
Best Local Similarity: 58.59% Mismatches: 33
Query Match: 60.70% Indels: 0

DB: 10 Gaps: 0

US-09-938-114-2 (1-129) x US-09-929-230-7 (1-580)

QY 1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
Db 36 GATTGTCCTCGGTGGTCCCTATGATGACGATGCTACAGGCTCATCAACAACTC 95
QY 21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
Db 96 AGGACGTGGGACGATGCGAGAGGTTCTGCTCGGAGCAGCGGAGCGGGCATCTCGTC 155
QY 41 SerLeuGluSerSerGlyGluAlaAspPheValGlyGlnLeuLeuAlaGlnLysLys 60
Db 156 TCTATGAAAGCGACGCGAAGCAGCCTTTGTGGCCAGCTGGTCCCTGAGAACATCAAG 215
QY 61 SerAlaLysLeuHisValTrpLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
Db 216 CAAAACAAATATGATGTCTGGATCGGACTGAGGATTCAAGGCGAAGAGCAATGCGAC 275
QY 81 IleGluTrpSerAspGlySerSerLeuSerLysGluAsnTrpIleGluGluSerLys 100
Db 276 ACGAAGTGGAGCGATGGCTCCAGCGCTCAATATGAGAACTGATTAAACATCGACCAA 335
QY 101 LysCysLeuGlyValHisLeuGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
Db 336 AAGTCTTTTGGGCTGAAAAAAGAGACAGAGGTTTCGACGCTGCGCAATGTTCACTGTACA 395
QY 121 GlnGlnAspPropheValCysGlu 128
Db 396 CAACAAATCTTTTCAATGTGCAAG 419

RESULT 5

US-10-226-420-7
Sequence 7, Application US/10226420
Publication No. US20030157686A1
GENERAL INFORMATION:
APPLICANT: Sheppard, Paul O.
APPLICANT: Bishop, Paul D.
TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
FILE REFERENCE: 00-72
CURRENT APPLICATION NUMBER: US/10/226,420
CURRENT FILING DATE: 2002-08-21
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 580
TYPE: DNA
ORGANISM: Sistrurus miliarius
FEATURE:
NAME/KEY: CDS
LOCATION: (3)...(434)
NAME/KEY: misc feature
LOCATION: (0)...(0)
OTHER INFORMATION: Zsnk4
US-10-226-420-7

Alignment Scores:

Pred. No.: 2.04e-51 Length: 580
Score: 434.00 Matches: 75
Percent Similarity: 74.22% Conservative: 20
Best Local Similarity: 58.59% Mismatches: 33
Query Match: 60.70% Indels: 0
DB: 13 Gaps: 0

US-09-938-114-2 (1-129) x US-10-226-420-7 (1-580)

QY 1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
Db 36 GATTGTCCTCGGTGGTCCCTATGATGACGATGCTACAGGCTCATCAACAACTC 95
QY 21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40


```

Db 190 WSNATHGARWSNWSNGARGCNGCNTTYTGNGCNCARYTNGTNCNCARAYMGNMGN 249
Qy 61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
Db 250 MNGMGNATHYTNATVATHTGGATHTGGTNGTNGTNCARGGNAARGAARCAARTGYWEN 309
Qy 81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluSerLys 100
Db 310 GCNAARTGGWSNGATGGNWSNWSNGTNTWNTAYGARAATYGGATHGARGCNGARWSNAAR 369
Qy 101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
Db 370 ACNTGYTNGGNYTNCARGCNGACNACNAATYATYCAATGGGTWATYATHTAYTYGNGN 429
Qy 121 GlnGlnAspPropheValCysGluAla 129
Db 430 GARATHAAYCNCNTTYGTNTGYGARGCN 456

RESULT 8
US-09-929-230-10
; Sequence 10, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/09/929,230
; CURRENT FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 725
; TYPE: DNA
; ORGANISM: Sistrurus miliarius
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (88)...(561)
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Zanks
US-09-929-230-10

Alignment Scores:
Pred. No.: 1,19e-44 Length: 725
Score: 387.00 Matches: 67
Percent Similarity: 71.88% Conservative: 25
Best Local Similarity: 52.34% Mismatches: 36
Query Match: 54.13% Indels: 0
DB: 10 Gaps: 0

US-09-938-114-2 (1-129) x US-09-929-230-10 (1-725)
Qy 1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
Db 163 AATTGTCCCTCTGTTGGTTCGCTACGATCAGTATTGCTACAGGTCATCAACGACTC 222
Qy 21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
Db 223 AAGACCTGGGACGATGACAGACGGTCTTCTGTCGAGCAGCGAGGCGGCGCATCTGGCG 282
Qy 41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
Db 283 TCTGTCAAAACGATGAGAACGAGTCTTCTGCCCAGTGTGGTTCGCGAACAATAAAG 342
Qy 61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
Db 343 CAAACCAATACTATGTCTGGATTTGACTGAGGATTCAAACAAAGACAGCAATGACGC 402
Qy 81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluSerLys 100
Db 403 ACGAAGTGGAGCGGATGCTCAGCGTTCAGTATGAGAACCTGGTTAAATCATTCCAAA 462
Qy 101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
Db 463 AAGTGTCTTGGCTGAAAAAAGACAGACAGTCTTCTCAATGGTACATACTACTGCTGCGAA 522
Qy 121 GlnGlnAspPropheValCysGlu 128
Db 523 GAAAAAAACCTTTTCTGCTGCAAG 546

```

```

Qy 101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
Db 463 AAGTGTCTTGGCTGAAAAAAGACAGACAGTCTTCTCAATGGTACATACTACTGCTGCGAA 522
Qy 121 GlnGlnAspPropheValCysGlu 128
Db 523 GAAAAAAACCTTTTCTGCTGCAAG 546

RESULT 9
US-10-226-420-10
; Sequence 10, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/10/226,420
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 725
; TYPE: DNA
; ORGANISM: Sistrurus miliarius
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (88)...(561)
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Zanks
US-10-226-420-10

Alignment Scores:
Pred. No.: 1,19e-44 Length: 725
Score: 387.00 Matches: 67
Percent Similarity: 71.88% Conservative: 25
Best Local Similarity: 52.34% Mismatches: 36
Query Match: 54.13% Indels: 0
DB: 13 Gaps: 0

US-09-938-114-2 (1-129) x US-10-226-420-10 (1-725)
Qy 1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
Db 163 AATTGTCCCTCTGTTGGTTCGCTACGATCAGTATTGCTACAGGTCATCAACGACTC 222
Qy 21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
Db 223 AAGACCTGGGACGATGACAGACGGTCTTCTGTCGAGCAGCGAGGCGGCGCATCTGGCG 282
Qy 41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
Db 283 TCTGTCAAAACGATGAGAACGAGTCTTCTGCCCAGTGTGGTTCGCGAACAATAAAG 342
Qy 61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
Db 343 CAAACCAATACTATGTCTGGATTTGACTGAGGATTCAAACAAAGACAGCAATGACGC 402
Qy 81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluSerLys 100
Db 403 ACGAAGTGGAGCGGATGCTCAGCGTTCAGTATGAGAACCTGGTTAAATCATTCCAAA 462
Qy 101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
Db 463 AAGTGTCTTGGCTGAAAAAAGACAGACAGTCTTCTCAATGGTACATACTACTGCTGCGAA 522
Qy 121 GlnGlnAspPropheValCysGlu 128
Db 523 GAAAAAAACCTTTTCTGCTGCAAG 546

```

RESULT 10

US-09-929-230-9
; Sequence 9, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/09/929,230
; CURRENT FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 432
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: This degenerate nucleotide sequence encodes the
; OTHER INFORMATION: amino acid sequence of SEQ ID NO:8.
; NAME/KEY: misc feature
; LOCATION: 6, 15, 18, 21, 24, 27, 42, 45, 57, 78, 81, 93, 96, 99, 111,
; LOCATION: 117, 126, 135, 141, 144, 150, 153, 156, 165, 171, 177, 180,
; LOCATION: 186, 189, 195, 198, 201, 231, 240, 243, 246, 255, 273, 276,
; LOCATION: 285, 291, 294, 297, 300, 315, 327, 330, 345, 348, 360
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: 363, 369, 372, 378, 384, 393, 405, 423, 426
; OTHER INFORMATION: n = A,T,C or G
US-09-929-230-9

Alignment Scores:

Pred. No.: 1.06e-43 Length: 432
Score: 378.00 Matches: 64
Percent Similarity: 62.50% Conservative: 16
Best Local Similarity: 50.00% Mismatches: 48
Query Match: 52.87% Indels: 0
DB: 10 Gaps: 0

US-09-938-114-2 (1-129) x US-09-929-230-9 (1-432)

QY 1 AspCysSerSerAspTTPSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
Db 34 GAYTGCCWNSGAYTGGTAYGCNTAYGAYCARTAYGTATMGNTNATHAARCAATYN 93
QY 21 LysThrTTPThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
Db 94 MGNACNTGGGAYGAYGCNGARMGNTTYTGYWNGARCARCNGAARGGNGCAYTNGTN 153
QY 41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
Db 154 WSNATHGARWSNGAYGNGGNGCNGTNTTYTGYWNGARCARCNGAARGGNGCAYTNGTN 213
QY 61 SerAlaLysIleHisValTTPileGlyLeuAgaGlnAsnLysGlnLysGlnCysSer 80
Db 214 CARAAAYAAATAYGAYTNTGGATHGNTNMGNTNATHCARGGNGARCARAARCAATYWSN 273
QY 81 IleGluTTPSerAspGlySerSerIleSerLysGluAsnTTPileGluGluSerLys 100
Db 274 ACNAAATGWSNGAYGNGWSNNGTNAAYTAYGARAAYTATNATHAARCAAYGNCNAAAR 333
QY 101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTTPcluanPheTyrCysGlu 120
Db 334 AARTGTYTGGNTNAAARARACNGGNTTYTGMNACNTGTMGNAAAYTNCAYTGYACN 393
QY 121 GlnGlnAspProPheValCysGlu 128
Db 394 CARCARAAAYTNTTATGTGYAAR 417

RESULT 11

US-10-226-420-9
; Sequence 9, Application US/10226420
; Publication No. US20030157686A1

GENERAL INFORMATION:

; APPLICANT: Sheppard, Paul O.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/10/226,420
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 432
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: This degenerate nucleotide sequence encodes the
; OTHER INFORMATION: amino acid sequence of SEQ ID NO:8.
; NAME/KEY: misc feature
; LOCATION: 6, 15, 18, 21, 24, 27, 42, 45, 57, 78, 81, 93, 96, 99, 111,
; LOCATION: 117, 126, 135, 141, 144, 150, 153, 156, 165, 171, 177, 180,
; LOCATION: 186, 189, 195, 198, 201, 231, 240, 243, 246, 255, 273, 276,
; LOCATION: 285, 291, 294, 297, 300, 315, 327, 330, 345, 348, 360
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: 363, 369, 372, 378, 384, 393, 405, 423, 426
; OTHER INFORMATION: n = A,T,C or G
US-10-226-420-9

Alignment Scores:

Pred. No.: 1.06e-43 Length: 432
Score: 378.00 Matches: 64
Percent Similarity: 62.50% Conservative: 16
Best Local Similarity: 50.00% Mismatches: 48
Query Match: 52.87% Indels: 0
DB: 13 Gaps: 0

US-09-938-114-2 (1-129) x US-10-226-420-9 (1-432)

QY 1 AspCysSerSerAspTTPSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
Db 34 GAYTGCCWNSGAYTGGTAYGCNTAYGAYCARTAYGTATMGNTNATHAARCAATYN 93
QY 21 LysThrTTPThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
Db 94 MGNACNTGGGAYGAYGCNGARMGNTTYTGYWNGARCARCNGAARGGNGCAYTNGTN 153
QY 41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
Db 154 WSNATHGARWSNGAYGNGGNGCNGTNTTYTGYWNGARCARCNGAARGGNGCAYTNGTN 213
QY 61 SerAlaLysIleHisValTTPileGlyLeuAgaGlnAsnLysGlnLysGlnCysSer 80
Db 214 CARAAAYAAATAYGAYTNTGGATHGNTNMGNTNATHCARGGNGARCARAARCAATYWSN 273
QY 81 IleGluTTPSerAspGlySerSerIleSerLysGluAsnTTPileGluGluSerLys 100
Db 274 ACNAAATGWSNGAYGNGWSNNGTNAAYTAYGARAAYTATNATHAARCAAYGNCNAAAR 333
QY 101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTTPcluanPheTyrCysGlu 120
Db 334 AARTGTYTGGNTNAAARARACNGGNTTYTGMNACNTGTMGNAAAYTNCAYTGYACN 393
QY 121 GlnGlnAspProPheValCysGlu 128
Db 394 CARCARAAAYTNTTATGTGYAAR 417

RESULT 12

US-09-929-230-12
; Sequence 12, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.


```

; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: RATTLESLAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/09/929,230
; CURRENT FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 474
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: This degenerate nucleotide sequence
; OTHER INFORMATION: amino acid sequence of SEQ ID NO
; NAME/KEY: misc_feature
; LOCATION: 6, 921, 24, 30, 33, 36, 39, 42, 48, 51,
; LOCATION: 66, 69, 84, 87, 90, 99, 120, 123, 132, 133,
; LOCATION: 168, 177, 183, 186, 192, 195, 198, 201, 212,
; LOCATION: 237, 240, 243, 246, 273, 282, 285, 288, 300
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc_feature
; LOCATION: 327, 339, 336, 339, 342, 345, 357, 360, 363
; LOCATION: 402, 411, 426, 447, 453, 465, 468
; OTHER INFORMATION: n = A,T,C or G
; US-09-929-230-12

```

Alignment Scores:		
Pred. No.:	9,818-37	474
Score:	329.00	55
Percent Similarity:	58.5%	20
Best Local Similarity:	42.97%	53
Query Match:	46.01%	20
DB:	10	0
		Gaps: 0
	Length:	Matches:
		Mismatches:
		Indels:
		Gaps:

US-09-938-114-2 (1-129) x US-09-929-230-12 (1-474)

[illegible]

RESULT 13

US-10-226-420-12
; Sequence 12, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72

```

; CURRENT APPLICATION NUMBER: US/10/226,420
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 474
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: This degenerate nucleotide sequence encodes the
; OTHER INFORMATION: amino acid sequence of SEQ ID NO:11.

```

```

? OTHER INFORMATION: This degenerate nucleotide sequence encodes the
? OTHER INFORMATION: amino acid sequence of SEQ ID NO:11.
? NAME/KEY: misc_feature
? LOCATION: 6, 9, 21, 24, 30, 33, 36, 39, 42, 48, 51, 54, 57, 60, 63,
? LOCATION: 66, 69, 84, 87, 90, 99, 120, 123, 132, 135, 141, 153, 159,
? LOCATION: 168, 177, 183, 186, 192, 195, 198, 201, 219, 222, 228, 231,
? LOCATION: 237, 240, 243, 246, 273, 282, 285, 288, 303, 315, 318,
? OTHER INFORMATION: n = A,T,C or G

```

```

; NAME/KEY: misc_feature
; LOCATION: 327, 339, 336, 339, 342, 345, 357, 360, 366, 372, 387, 390,
; LOCATION: 402, 411, 426, 447, 453, 465, 468
; OTHER INFORMATION: n = A,T,C or G
US-09-929-230-12

```

Alignment Scores:
Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:

US-09-938-114-2 (1-129) x US-10-226-420-12 (1-474)

1	AspCysSerSerAspTrpSerSerIyGluClyHisCysTyrIysValPheIysGlnSer	20
YY		
76	AAYTGCCNWSNGNGTGGTYGCTAYGAYCARTAYGTGYAYMGNGTNAATHARMGNYTN	135
YY		
21	LysThrTrpThrAspAlaGluSerPheCysThrIysGlnValAsnGlyGlyHisLeuVal	40
YY		
136	AARACNTGGGAYGAYGCGNARMGNVTITGYWSNGARCARGCNAARGGNGNCAYTTNGCN	195
YY		
41	SerLeGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnIysIleLys	60
YY		
196	WSNGTNGARAAYGAYGARGGCGNTNTTYTNGCNCARYTNGTNGCNGCNAAYATHAAR	255
YY		
61	SerAlaIysIleHisValTrpIleGlyLeuArgAlaGlnAsnIysGluIysGlnCysSer	80
YY		
256	CARAAYCARTAYTAYGNTGTGATGGTYTNGMGNATHCARAAVAARGCNCARCARTGYWSN	315
YY		
81	IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerIys	100
YY		
316	ACNAARPTGGWSNGAYGWNWSNGTNGWSNTAYGARAAYTTGTNAARWSNCAYWSNAAR	375
YY		
101	LysCysLeuGlyValHisIleGluThrGlyPheHisIlyTrpGluAsnPheTyrCysGlu	120
YY		
376	AARTGYTTYGNTYNAARAARGARACNGARTTYTNCARTGGTAYAAACNGAYTGYGAR	435
YY		
121	GlnGlnAspProPheValCysGlu	128
YY		
436	GARAARAAYTNTTYGTTNTGYAAR	459
YY		

RESULT 14

US-09-969-763-2
; Sequence 2, Application US/09896763
; Publication No. US20020198363A1
; GENERAL INFORMATION:
; APPLICANT: FUKUCHI, NAOYUKI
; APPLICANT: KITO, MORIKAZU
; APPLICANT: KAYAHARA, TAKASHI
; APPLICANT: FUTAKI, FUMIE
; APPLICANT: ISHIKAWA, KOHKI

```
; APPLICANT: SUZUKI, EIICHIRO
; APPLICANT: GONDOH, KEIKO
; APPLICANT: SHIMBA, NOBUHISA
; APPLICANT: YAMADA, NAOKYUKI
; TITLE OF INVENTION: PROTEIN HAVING ANTITHROMBOTIC ACTIVITY AND METHOD FOR PRODUCING IT
; FILE REFERENCE: 214760USO
; CURRENT APPLICATION NUMBER: US/09/969,763
; CURRENT FILING DATE: 2000-10-25
; PRIOR APPLICATION NUMBER: JP 2000-305279
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 690
; TYPE: DNA
; ORGANISM: Crotalus harridus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (66)..(512)
; OTHER INFORMATION:
US-09-969-763-2

Alignment Scores:
Pred. No.: 1,658-30 Length: 690
Score: 286.50 Matches: 55
Percent Similarity: 60.16% Conservative: 22
Best Local Similarity: 42.97% Mismatches: 42
Query Match: 40.07% Indels: 9
DB: 10 Gaps: 2

US-09-938-114-2 (1-129) x US-09-969-763-2 (1-690)
QY 1 AspCysSerSerAspTyrSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
Db 141 GAATGTCCTCGGTTGGTCTTCATGATCGTATGCTACAGCCCTTCAACAAGAG 200
QY 21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
Db 201 ATGACTGGCGCGATGCGAGAGGTTCTCTCGAGACGGCGAAGGGCGGCGATCTCTC 260
QY 41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
Db 261 TCTGCGAAACCGCCCTAGAACATCTTTTGGCAATGTCTCTATCGGAACAAGAG 320
QY 61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGlnLysCysSer 80
Db 321 TACCTCACACGTTATATCTGATTCGATTCGAGGTTTCAAAACAAGAGCAGCCATGC 377
QY 81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100
Db 378 -----TCCAGCATCATGAGAACCTGTT-----GACCCATTT 413
QY 101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
Db 414 GAATGTTTTATGTTGAGACAGACACAAAGCTCGTGAGTGGTTTAAAGTTGACTGTGAA 473
QY 121 GlnGlnAspPropheValCysGlu 128
Db 474 CAACACATTTCTTCATATGCAAG 497

RESULT 15
US-09-929-230-1
; Sequence 1, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: RATTLESLAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/09/929,230
; CURRENT FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 1
; LENGTH: 483
; TYPE: DNA
; ORGANISM: Sistrurus miliarius
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(455)
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Zank2
US-09-929-230-1

Alignment Scores:
Pred. No.: 1,48e-27 Length: 483
Score: 264.00 Matches: 55
Percent Similarity: 57.03% Conservative: 18
Best Local Similarity: 42.97% Mismatches: 49
Query Match: 36.92% Indels: 6
DB: 10 Gaps: 3

US-09-938-114-2 (1-129) x US-09-929-230-1 (1-483)
QY 1 AspCysSerSerAspTyrSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
Db 60 GATTGTCCTCTGACTGCTCTCTCTATGATCAGCATTCGTACAGGTCTTCAGTGAATC 119
QY 21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
Db 120 AAAACCTGGGATGATGCAGAGAGTTTCTGCTACACACAGACAGAGCCGCTGGCC 179
QY 41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
Db 180 TCCATCCACAGCAGTGAAGAGAGACTTTTGTGGCAAACTGGCCTCCCAACTTTGAAA 239
QY 61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGlnLysCysSer 80
Db 240 TTCACCTTC-----ATGTGGATCGACTGAAA-----GATCTATGGAAGAATGCAAA 287
QY 81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100
Db 288 TGGCAGTGGAGCGATGACACCAAACTGGACTACAAAGCCTGGACTCGAAGA-----CCC 341
QY 101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
Db 342 TATTGTACAGTAATGTTAGTCAAGACAGATAGGATCTTTTGGTTCAATAGAGTTGCGAA 401
QY 121 GlnGlnAspPropheValCysGlu 128
Db 402 AAGACTGTATCTTTTGTCTGCAAG 425

Search completed: December 8, 2003, 19:36:04
Job time : 516.051 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:49:54 ; Search time 25.8 Seconds
(without alignments)
480.843 Million cell updates/sec

Title: US-09-938-114-2

Perfect score: 715

Sequence: 1 DCSSDWSYEGHCYKFKQS.....GFHKWENFYCQQDPFVCEA 129

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	685	95.8	152	2 JC71134	agkisacutacin alph
2	527	73.7	129	2 JC4329	coagulation factor
3	496	69.4	152	2 JC4690	coagulation factor
4	422.5	59.1	133	2 A47267	botrocetin alpha c
5	420.5	58.8	144	2 PC7027	aggreitin alpha cha
6	345	48.3	131	2 JC5058	bitiscetin alpha c
7	291	40.7	146	2 JC4691	coagulation factor
8	280	39.2	146	2 JC7105	aggreitin beta chai
9	274	38.3	146	2 JC7135	agkisacutacin beta
10	271	37.9	125	2 JC5059	bitiscetin beta ch
11	260	36.4	125	2 B47267	botrocetin beta ch
12	252	35.2	123	2 JC2415	echicetin beta cha
13	248	34.7	123	2 B42972	coagulation factor
14	231.5	32.4	135	2 A38609	lectin, galactose-
15	201.5	28.2	166	1 RGHUIA	regenerating islet
16	201.5	28.2	166	2 A45751	pancreatic stone p
17	199.5	27.9	1340	2 A39808	proteoglycan core
18	199.5	27.9	2327	2 T42630	aggrean - bovine
19	197.5	27.6	1479	2 T42710	mannose receptor,
20	194	27.1	172	2 S32489	lectin - Iberian r
21	192.5	26.9	166	1 RGHUIB	regenerating islet
22	190.5	26.6	165	2 A47148	reg I, regeneratin
23	189.5	26.5	173	2 B47148	reg II, regenerati
24	189.5	26.5	2124	2 A28452	proteoglycan core
25	189.5	26.5	2132	1 A55182	aggrean precursor
26	186.5	26.1	174	2 S54979	pancreatitiss-assoc
27	186	26.0	175	2 A37194	pancreatic thread
28	186	26.0	175	2 S29822	pancreatitiss-assoc
29	185.5	25.9	1268	2 S52781	neurocan - mouse

ALIGNMENTS

RESULT 1

JC71134

agkisacutacin alpha chain precursor - sharp-nosed viper

N:Alternate names: fibrinogenolytic venom protein

C:Species: Agkistrodon acutus (sharp-nosed viper)

C>Date: 04-Mar-2000 #sequence_revision 04-Mar-2000 #text_change 24-Oct-2000

C:Accession: JC71134; PC7037

R:Cheng, X.; Qian, Y.; Liu, Q.; Li, B.X.Y.; Zhang, M.; Liu, J.

Biochem. Biophys. Res. Commun. 265, 530-535, 1999

A:Title: Purification, characterization, and cDNA cloning of a new fibrinogenolytic venom

A:Reference number: JC71134; MUID:20025379; PMID:10558903

A:Accession: JC71134

A:Molecule type: mRNA

A:Residues: 1-152 <CH>

A:Cross-references: GB:AF176420

A:Experimental source: venom gland

A:Accession: PC7037

A:Molecule type: protein

A:Residues: 24-53;84-86;87-94;125-136;137-152 <CH2>

C:Superfamily: tetranectin; C-type lectin homology

C:Keywords: disulfide bond; heterodimer; venom

F:1-23/Domain: signal sequence #status predicted <SIG>

F:24-152/Product: agkisacutacin alpha chain #status experimental <MAT>

Query Match

Best Local Similarity 95.8%; Score 685; DB 2; Length 152;

Matches 124; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 DCSSDWSYEGHCYKFKQSKTWTDAESFCTKQVNGHLVSISSGEADFYGLIAQKIK 60

Db 24 DCSSGWSYEGHCYKFKQSKTWADAEFCTKQVNGHLVSISSGEADFYGLIAQKIK 83

Qy 61 SAKIHWIGLRAQNKQCKSIENWSDGSSISKENWIEESKCLGVHETGFHKWENFYCE 120

Db 84 SAKIHWIGLRAQNKQCKSIENWSDGSSISKENWIEESKCLGVHETGFHKWENFYCE 143

Qy 121 QQDPFVCEA 129

Db 144 QQDPFVCEA 152

RESULT 2

JC4329

coagulation factor IX-binding protein A chain - habu

C:Species: Trimeresurus flavoviridis (habu)

C>Date: 06-Dec-1995 #sequence_revision 08-Feb-1996 #text_change 13-Mar-1998

C:Accession: JC4329

R:Atoda, H.; Ishikawa, M.; Yoshihara, E.; Sekiya, F.; Morita, T.

J. Biochem. 118, 965-973, 1995

A:Title: Blood coagulation factor IX-binding protein from the venom of Trimeresurus flav

A:Reference number: JC4329; MUID:96318509; PMID:8749314

A:Accession: JC4329

```
RESULT 4
A47267
Molecule type: protein
A;Residues: 1-129 <ATO>
C;Comment: This protein binds calcium.
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: anticoagulant; blood coagulation; calcium binding; venom
F;2-127/Domain: C-type lectin homology <LCH>
F;2-13,30-127,102-113/Disulfide bonds: #status predicted
Query Match 73.7%; Score 527; DB 2; Length 129;
Best Local Similarity 72.1%; Pred. No. 9e-44;
Matches 93; Conservative 10; Mismatches 26; Indels 0; Gaps 0;
QY 1 DCSSDWSYEGHCYKVFQKQSTWTDASFCYKQVNGHLSVIESGGEADFGVQLIAQKIK 60
Db 1 DCPGWSYEGHCYKVFQKQSTWTDASFCYKQVNGHLSVIESGGEADFGVQLIAQKIK 60
QY 61 SAKIHWIGLRAQNKKEKCSIEWSDGSSISKENWIEESKCLGVHIEFGFKWENFYCE 120
Db 61 NTKSYVWIGLRAQNKKEKCSIEWSDGSSISKENWIEESKCLGVHIEFGFKWENFYCE 120
QY 121 QDDPFVCEA 129
Db 121 QQNPFVCEA 129
RESULT 3
JC4690
coagulation factor IX/factor X-binding protein chain A precursor - habu
C;Species: Trimeresurus flavoviridis (habu)
C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 20-Jun-2000
C;Accession: JC4690; A39332
R;Matsuzaki, R.; Yoshiara, E.; Yamada, M.; Shima, K.; Atoda, H.; Morita, T.
Biochem. Biophys. Res. Commun. 220, 382-387, 1996
A;Title: cDNA cloning of IX/X-BP, a heterogeneous two-chain anticoagulant protein from s
A;Reference number: JC4690; MUID:96184662; PMID:8645314
A;Accession: JC4690
A;Molecule type: mRNA
A;Residues: 1-152 <MATL>
A;Cross-references: DDBJ:D83331; NID:gl402639; PIDN:BAAL1887.1; PID:gl402640
A;Experimental source: venom
R;Atoda, H.; Hyuga, M.; Morita, T.
J. Biol. Chem. 266, 14903-14911, 1991
A;Title: The primary structure of coagulation factor IX/factor X-binding protein isolate
oetin, tetranectin, and lymphocyte Fc epsilon receptor for immunoglobulin E.
A;Reference number: A39332; MUID:91332000; PMID:1831197
A;Accession: A39332
A;Status: preliminary
A;Molecule type: protein
A;Residues: 24-152 <ATO>
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: anticoagulant; blood coagulation; glycoprotein; hemolymph; lectin
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-152/Product: factor IX/X binding protein chain A #status predicted <MAT>
F;25-150/Domain: C-type lectin homology <LCH>
F;25-36,53-150,125-142/Disulfide bonds: #status predicted
Query Match 69.4%; Score 496; DB 2; Length 152;
Best Local Similarity 67.4%; Pred. No. 1.1e-40;
Matches 87; Conservative 11; Mismatches 31; Indels 0; Gaps 0;
QY 1 DCSSDWSYEGHCYKVFQKQSTWTDASFCYKQVNGHLSVIESGGEADFGVQLIAQKIK 60
Db 24 DCLSGWSYEGHCYKVFQKQSTWTDASFCYKQVNGHLSVIESGGEADFGVQLIAQKIK 83
QY 61 SAKIHWIGLRAQNKKEKCSIEWSDGSSISKENWIEESKCLGVHIEFGFKWENFYCE 120
Db 84 RLDFYIWIGLRVGKVKQCNSEWSGSSVYENWIEAESKTCGLGLEKETDFRKWNVYCG 143
QY 121 QDDPFVCEA 129
Db 144 QQNPFVCEA 152
```

```
RESULT 4
A47267
botrocetin alpha chain - jararaca
N;Alternate names: two chain botrocetin alpha chain
C;Species: Bothrops jararaca (jararaca)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 15-May-1998
C;Accession: A47267; B37958
R;Usami, Y.; Fujimura, Y.; Suzuki, M.; Ozeki, Y.; Nishio, K.; Fukui, H.; Titani, K.
Proc. Natl. Acad. Sci. U.S.A. 90, 928-932, 1993
A;Title: Primary structure of two-chain botrocetin, a von Willebrand factor modulator pu
A;Reference number: A47267; MUID:93157385; PMID:8430107
A;Accession: A47267
A;Molecule type: protein
A;Residues: 1-133 <USA>
A;Experimental source: venom
A;Note: sequence extracted from NCBI backbone (NCBI:124085)
R;Fujimura, Y.; Titani, K.; Usami, Y.; Suzuki, M.; Oyama, R.; Matsui, T.; Fukui, H.; Sug
Biochemistry 30, 1957-1964, 1991
A;Title: Isolation and chemical characterization of two structurally and functionally di
A;Reference number: A37958; MUID:91129280; PMID:1993206
A;Accession: B37958
A;Molecule type: protein
A;Residues: 1-40 <FUG>
C;Complex: heterodimer of alpha and beta (see PIR:B47267) chains
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: hemagglutinin; heterodimer; venom
F;2-128/Domain: C-type lectin homology <LCH>
F;2-13,30-128,103-120/Disulfide bonds: #status experimental
F;80/Disulfide bonds: interchain (to beta-75) #status experimental
Query Match 59.1%; Score 422.5; DB 2; Length 133;
Best Local Similarity 56.9%; Pred. No. 1.1e-33;
Matches 74; Conservative 20; Mismatches 35; Indels 1; Gaps 1;
QY 1 DCSSDWSYEGHCYKVFQKQSTWTDASFCYKQVNGHLSVIESGGEADFGVQLIAQKIK 59
Db 1 DCPGWSYEGHCYKVFQKQSTWTDASFCYKQVNGHLSVIESGGEADFGVQLIAQKIK 60
QY 60 KSAKHWIGLRAQNKKEKCSIEWSDGSSISKENWIEESKCLGVHIEFGFKWENFYC 119
Db 61 QSSDLIAWIGLRAQNKKEKCSIEWSDGSSISKENWIEESKCLGVHIEFGFKWENFYC 120
QY 120 EQDDPFVCEA 129
Db 121 AQKPFVCKS 130
RESULT 5
PC7027
aggrexin alpha chain - Malayan pit viper (fragment)
C;Species: Calloselasma rhodostoma (Malayan pit viper)
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 02-Jun-2000
C;Accession: PC7027
R;Chung, C.H.; Au, L.C.; Huang, T.F.
Biochem. Biophys. Res. Commun. 263, 723-727, 1999
A;Title: Molecular cloning and sequence analysis of aggrexin, a collagen-like platelet
A;Reference number: PC7027; MUID:99443731; PMID:10512747
A;Accession: PC7027
A;Molecule type: mRNA
A;Residues: 1-144 <CHU>
A;Experimental source: venom gland
A;Superfamily: tetranectin; C-type lectin homology
C;Keywords: disulfide bond; platelet aggregation; venom
Query Match 58.8%; Score 420.5; DB 2; Length 144;
Best Local Similarity 59.7%; Pred. No. 1.9e-33;
Matches 77; Conservative 16; Mismatches 35; Indels 1; Gaps 1;
QY 1 DCSSDWSYEGHCYKVFQKQSTWTDASFCYKQVNGHLSVIESGGEADFGVQLIAQKIK 60
Db 12 DCDFGNSPYDQHCYQAFNEQKTDWEAEKFCRAQGAHLASIESGGEADFGVQLIAQKIK 71
QY 61 SAKIHWIGLRAQNKKEKCSIEWSDGSSISKENWIEESKCLGVHIEFGFKWENFYC 119
```

A11:The primary structure of coagulation factor IX/factor X-binding protein isolate
J. Biol. Chem. 266, 14903-14911, 1991

Db 64 --VLVWIGL--SHFWRCPLRWTDGARDLYRALSDEPI--CF---VAESFNKWIQWTCN 114
Qy 121 QQDPFVCE 128
Db 115 RKKSFVCK 122

RESULT 11
B47267
botrocetin beta chain - jararaca
N;Alternate names: two chain botrocetin beta chain
C;Species: Bothrops jararaca (jararaca)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 15-May-1998
C;Accession: B47267; C37958
R;Usami, Y.; Fujimura, Y.; Suzuki, M.; Ozeki, Y.; Nishio, K.; Fukui, H.; Titani
Proc. Natl. Acad. Sci. U.S.A. 90, 928-932, 1993
A;Title: Primary structure of two-chain botrocetin, a von Willebrand factor mod
A;Reference number: A47267; MUID:93157385; PMID:8430107
A;Accession: B47267
A;Molecule type: protein
A;Residues: 1-125 <USA>
A;Experimental source: venom
A;Note: sequence extracted from NCBI backbone (NCBIP:124086)
R;Fujimura, Y.; Titani, K.; Usami, Y.; Suzuki, M.; Oyama, R.; Matsui, T.; Fukui
Biochemistry 30, 1957-1964, 1991
A;Title: Isolation and chemical characterization of two structurally and functi
A;Reference number: A37958; MUID:91129280; PMID:1993206
A;Accession: C37958
A;Molecule type: protein
A;Residues: 1-40 <PU>
C;Complex: heterodimer of alpha (see PIR:A47267) and beta chains
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: hemagglutinin; heterodimer; venom
E;2-121/Domain: C-type lectin homology,<LCH>
F;2-13,30-121,98-113/Disulfide bonds: #status experimental
F;75/Disulfide bonds: interchain (to alpha-80) #status experimental

Qy 1 DCSDWSSEYGHCHVKFKQKWTDAESFCTQVNGHLSVIESGSEADFGQLIAQIK 60
Db 1 DCPDPWSSEYGHCHVKFKQKWTDAESFCTQVNGHLSVIESGSEADFGQLIAQIK 60
Qy 61 SAKTHWIGLRANQKEQCSLEWSDGSSISKENWIEESKKCLGVIHTGTGFKWENFYCE 120
Db 61 GDV--VWIGL--SDVWVKCFEWDGMEFDYDYLLIAEYEC--VASKPTNKKWIIPECT 114
Qy 121 QQDPFVCE 128
Db 115 RFXNFVCE 122

RESULT 12
JC2415
echicetin beta chain - saw-scaled viper
C;Species: Echis carinatus (saw-scaled viper)
C;Date: 21-Mar-1995 #sequence_revision 26-May-1995 #text_change 15-Oct-1996
C;Accession: JC2415
R;Peng, M.; Holt, J.C.; Niewiarowski, S.
Biochem. Biophys. Res. Commun. 205, 68-72, 1994
A;Title: Isolation, characterization and amino acid sequence of echicetin beta
A;Reference number: JC2415; MUID:95091801; PMID:7999097
A;Accession: JC2415
A;Molecule type: protein
A;Residues: 1-123 <PEN>
A;Experimental source: venom
A;Comment: This protein inhibits agglutination of fixed platelets induced by s
clebrand factor and alboaggregins.
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: anticoagulant; dimer

F:2-119/Domain: C-type lectin homology <LCH>
F:2-13,30-119,96-111/Disulfide bonds: #status predicted

Query Match 35.2%; Score 252; DB 2; Length 123;
Best Local Similarity 40.6%; Pred. No. 3e-17;
Matches 52; Conservative 18; Mismatches 50; Indels 8; Gaps 4;

QY 1 DCSSDWSSEYGHCVKFKQSKTWTDAESFCTKVQNGHLVSISSGEADFGQLIAQKIK 60
DB 1 NCLPDMSSVYEGYCYKFKERNWADAEFCKQVKDGHLSVSPNSKEVDFMISLAPFMLK 60
QY 61 SAKIHVWIGLRAQNKKEKQCSIEWSGSSISKENWIEESKKCLGVHETGFKWENFYCE 120
DB 61 MEL--VWIGL--SDYWRDCYEWSDGALDYKAW--DNERHCFAA--KTIDNQWRRKGS 112
QY 121 QQDPFVCE 128
DB 113 GEYFVCK 120

RESULT 13
B42972
coagulation factor X activating enzyme (EC 3.4.24.-) light chain - Russell's viper
C/Species: Vipera russelli (Russell's viper)
C/Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 28-Feb-1997
C/Accession: B42972
F:/Takeya, H.; Nishida, S.; Miyata, T.; Kawada, S.; Saisaka, Y.; Morita, T.; Iwanaga, S.
J. Biol. Chem. 267, 14109-14117, 1992
A/Title: Coagulation factor X activating enzyme from Russell's viper venom (RVV-X). A no
A/Reference number: A42972; MUID:92332516; PMID:1629211
A/Contents: V. r. siamensis
A/Accession: B42972
A/Status: preliminary
A/Molecule type: protein
A/Residues: 1-123 <TAK>
A/Experimental source: venom
A/Note: sequence extracted from NCBI backbone (NCBIP:108408)
C/Superfamily: tetranectin; C-type lectin homology
C/Keywords: hydrolase; metalloproteinase; venom; zinc
F:4-121/Domain: C-type lectin homology <LCH>
F:4-15,32-121,98-113/Disulfide bonds: #status predicted

Query Match 34.7%; Score 248; DB 2; Length 123;
Best Local Similarity 43.8%; Pred. No. 7.3e-17;
Matches 56; Conservative 16; Mismatches 48; Indels 8; Gaps 5;

QY 1 DCSSDWSSEYGHCVKFKQSKTWTDAESFCTKVQNGHLVSISSGEADFGQLIAQKIK 60
DB 3 DCPGSLVSEYGHCVKFGFNDLKNWTDKAEKFTQKKGHLVSLHGREEEFVNLISENLE 62
QY 61 SAKIHVWIGLRAQNKKEKQCSIEWSGSSISKENWIEESKKCLGVHETGFKWENFYCE 120
DB 63 YPA--TWIGL--GNWVKDCMEWSDRGNV-KYKALAES-YCL--IMTHEKEWMSMTCN 114
QY 121 QQDPFVCE 128
DB 115 FIAPVCK 122

RESULT 14
A38609
lectin, galactose-specific - western diamondback rattlesnake
C/Species: Crotalus atrox (western diamondback rattlesnake)
C/Date: 23-Aug-1991 #sequence_revision 23-Aug-1991 #text_change 11-Aug-1995
C/Accession: A38609
F:/Hirabayashi, J.; Kusunoki, T.; Kasai, K.
J. Biol. Chem. 266, 2320-2326, 1991
A/Title: Complete primary structure of a galactose-specific lectin from the venom of the
A/Reference number: A38609; MUID:91115849; PMID:1989986
A/Accession: A38609
A/Status: preliminary
A/Molecule type: protein
A/Residues: 1-135 <HIR>

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:46:14 ; Search time 16.9543 Seconds
(without alignments)
357.812 Million cell updates/sec

Title: US-09-938-114-2

Perfect score: 715

Sequence: 1 DCSSDWSSEYEGHYKVKQS.....GFHKWENFYCEQDDPFVCEA 129

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	496	69.4	152	1	IXA_TRIFL	P23806 trimeresuru
2	425	59.4	131	1	ABAI_TRIAB	P81111 trimeresuru
3	422.5	59.1	133	1	BOTA_BOTUA	P22029 bothrops ja
4	408	57.1	134	1	ABAB_TRIAB	P81112 trimeresuru
5	388	54.3	132	1	ABBA_TRIAB	P81115 trimeresuru
6	364	50.9	158	1	CVXA_CRODU	O93426 crotalus du
7	309	43.2	133	1	ECHA_ECHCA	P81017 echis carin
8	302	42.2	133	1	RHCA_AGRKH	P81397 agkistrodon
9	299	41.8	123	1	ABAA_TRIAB	P81114 trimeresuru
10	291	40.7	146	1	IXB_TRIFL	P23807 trimeresuru
11	273.5	38.3	127	1	CHBA_CROHO	P81508 crotalus ho
12	269	37.6	129	1	RHCB_AGRKH	P81398 agkistrodon
13	264	36.9	148	1	CVXB_CRODU	O93427 crotalus du
14	260	36.4	125	1	BOTB_BOTUA	P22030 bothrops ja
15	259	36.2	117	1	CHBB_CROHO	P81509 crotalus ho
16	252	35.2	123	1	ECHB_ECHCA	P81996 echis carin
17	235	32.9	118	1	ABBB_TRIAB	P81116 trimeresuru
18	231.5	32.4	135	1	LECG_CROAT	P21963 crotalus at
19	230	32.2	125	1	ABAB_TRIAB	P81113 trimeresuru
20	215.5	30.1	158	1	LECG_TRIST	O9YGP1 trimeresuru
21	209	29.2	175	1	PAP2_MOUSE	O09037 mus musculu
22	201.5	28.2	166	1	LITA_HUMAN	P05451 homo sapien
23	199.5	27.9	2364	1	PGCA_BOVIN	P13608 bos taurus
24	199	27.8	132	1	ACAL_ANSAN	P83308 anser anser
25	194.5	27.2	174	1	PAP3_MOUSE	O09049 mus musculu
26	194	27.1	172	1	LECA_PLEWA	O02988 pleurodeles
27	192.5	26.9	166	1	LITB_HUMAN	P48304 homo sapien
28	190.5	26.6	165	1	LIT1_MOUSE	P43137 mus musculu
29	189.5	26.5	173	1	LIT2_MOUSE	O08731 mus musculu
30	189.5	26.5	2124	1	PGCA_RAT	P07897 rattus norv
31	189.5	26.5	2132	1	PGCA_MOUSE	O61282 mus musculu
32	186.5	26.1	174	1	PAP3_RAT	P42854 rattus norv
33	186	26.0	175	1	LITH_BOVIN	P23132 bos taurus

34	186	26.0	175	1	PAP1_MOUSE	P35230 mus musculu
35	185.5	25.9	1268	1	PGCN_MOUSE	P5066 mus musculu
36	185.5	25.9	2333	1	PGCA_CANFA	O28343 canis fami
37	184.5	25.8	1357	1	PGCN_RAT	P5067 rattus norv
38	184.5	25.8	1321	1	PGCN_HUMAN	O14594 homo sapien
39	183.5	25.7	2109	1	PGCA_CHICK	P07898 gallus gall
40	183.5	25.7	2415	1	PGCA_HUMAN	P16112 homo sapien
41	183	25.6	175	1	PAP1_HUMAN	O06141 homo sapien
42	182.5	25.5	165	1	LITH_RAT	P10758 rattus norv
43	177	24.8	174	1	PAP2_RAT	P35331 rattus norv
44	175	24.5	175	1	PBGC_HUMAN	O92778 homo sapien
45	171	23.9	1458	1	PAP2_RABIT	P49260 oryctolagus

ALIGNMENTS

```

RESULT 1
IXA_TRIFL
ID IXA_TRIFL STANDARD; PRT; 152 AA.
AC P23806; Q91246;
DT 01-NOV-1991 (Rel. 20, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Coagulation factor IX/factor X-binding protein A chain precursor
DE (IX/X-BP).
OS Trimeresurus flavoviridis (Habu).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Trimeresurus.
OX NCBI_TaxID=88087;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96184662; PubMed=8645314;
RA Matsuzaki R., Yoshihara E., Yamada M., Shima K., Atoda H., Morita T.;
RT "cDNA cloning of IX/X-BP, a heterogeneous two-chain anticoagulant
RT protein from snake venom.";
RL Biochem. Biophys. Res. Commun. 220:382-387(1996).
RN [2]
RP SEQUENCE OF 24-152.
RC TISSUE=Venom;
RX MEDLINE=91332000; PubMed=1831197;
RA Atoda H., Hyuga M., Morita T.;
RT "The primary structure of coagulation factor IX/factor X-binding
RT protein isolated from the venom of Trimeresurus flavoviridis.
RT Homology with asialoglycoprotein receptors, proteoglycan core
RT protein, tetranectin, and lymphocyte Fc epsilon receptor for
RT immunoglobulin E.";
RL J. Biol. Chem. 266:14903-14911(1991).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX MEDLINE=97331317; PubMed=9187649;
RA Mizuno H., Fujimoto Z., Koizumi M., Kano H., Atoda H., Morita T.;
RT "Structure of coagulation factors IX/X-binding protein, a heterodimer
RT of C-type lectin domains.";
RL Nat. Struct. Biol. 4:438-441(1997).
CC -!- FUNCTION: ANTICOAGULANT PROTEIN WHICH BINDS WITH FACTOR IX AND
CC FACTOR X IN THE PRESENCE OF CALCIUM WITH A 1 TO 1 STOICHIOMETRY.
CC -!- SUBUNIT: HETERODIMER OF CHAINS A AND B LINKED BY A DISULFIDE BOND.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- MISCELLANEOUS: CALCIUM IS REQUIRED FOR LIGAND BINDING.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D83331; BAA11887.1; -
DR FIR; JC4690; JC4690.

```

```

DR PDB: 1LXX; 06-MAY-98.
DR InterPro; IPR001304; Lectin C.
DR Pfam; PF00059; Lactin c.1.
DR PRINTS; PR01504; PNCREATITAP.
DR SMART; SM00034; CLECT; 1. LECTIN 1; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS00041; C_TYPE_LECTIN_2; 1.
KW Lactin; Calcium; Signal; 3D-structure.
FT SIGNAL 1 23
FT CHAIN 24 152
FT DOMAIN 24 152
FT DISULFID 25 36
FT DISULFID 53 150
FT DISULFID 102 102
FT DISULFID 125 142
FT TURN 27 28
FT STRAND 30 31
FT STRAND 36 44
FT HELIX 46 56
FT TURN 58 59
FT STRAND 61 62
FT HELIX 68 81
FT STRAND 89 95
FT STRAND 105 105
FT TURN 107 108
FT STRAND 111 111
FT STRAND 117 117
FT HELIX 119 121
FT STRAND 125 128
FT HELIX 130 132
FT TURN 133 134
FT STRAND 137 140
FT TURN 142 143
FT STRAND 146 152
SQ SEQUENCE 152 AA; 17213 MW; FB3DDDD2369009263 CRC64;

Query Match 69.4%; Score 496; DB 1; Length 152;
Best Local Similarity 67.4%; Pred. No. 2.8e-40;
Matches 87; Conservative 11; Mismatches 31; Indels 0; Gaps 0;

QY 1 DCSWSSYEGHCYKVPKQSKTWTDAESFCTKVNGHVLVSIESGSEADVFVGLIAQKIK 60
Db 24 DCLSGWSSYEGHCYKAFKRYKTWEDAEKRVCTEQAGHVLVSIESGSEADVFVGLVQNWK 83

QY 61 SAKIHVWIGLRAQNKQKCSIEWSDGSSISKENWIEESKCLGVHIEFGFKWENFYCE 120
Db 84 RLDFYIWIGLVRVQGVKQCNSEWSDGSSSYENWIEESKTCGLEKETDFRKWNVIYCG 143

QY 121 QDDPFVCEA 129
Db 144 QQNPFVCEA 152

RESULT 2
ABAL TRIAB STANDARD; PRT; 131 AA.
AC P81111;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Alboaggregin A subunit 1.
OS Trimeresurus albolabris (White-lipped pit viper).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidodactylus; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Trimeresurus.
OX NCBI_TaxID=8765;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=98189535; PubMed=9531050;
RA Kowalska M.A., Tan L., Holt J.C., Peng M., Karczewski J.,

```

```

RA Calvete J.J., Niewiarowski S.;
RT "Alboaggregins A and B. Structure and interaction with human platelets.";
RL Thromb. Haemost. 79:609-613(1998).
CC -!- FUNCTION: Binds to platelet GPIb/IX receptor system and stimulates agglutination.
CC -!- SUBUNIT: Heterotetramer of the subunits 1, 2, 3 and 4, disulfide-linked.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
DR HSSP; P23806; 1LXX.
DR InterPro; IPR001304; Lectin C.
DR Pfam; PF00059; Lactin c.1.
DR SMART; SM00034; CLECT; 1. LECTIN 1; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS00041; C_TYPE_LECTIN_2; 1.
KW Lactin.
FT DOMAIN 1 129 C-TYPE LECTIN (LONG FORM).
FT DISULFID 2 13 BY SIMILARITY.
FT DISULFID 30 127 BY SIMILARITY.
FT DISULFID 102 119 BY SIMILARITY.
SQ SEQUENCE 131 AA; 15427 MW; B3569F5BF91F6624 CRC64;

Query Match 59.4%; Score 425; DB 1; Length 131;
Best Local Similarity 57.0%; Pred. No. 1.3e-33;
Matches 73; Conservative 22; Mismatches 33; Indels 0; Gaps 0;

QY 1 DCSWSSYEGHCYKVPKQSKTWTDAESFCTKVNGHVLVSIESGSEADVFVGLIAQKIK 60
Db 1 DCPDWSYDQCYRVFRKIQTWEDAEKRVCTEQAGHVLVSIESGSEADVFVGLVSENIR 60

QY 61 SAKIHVWIGLRAQNKQKCSIEWSDGSSISKENWIEESKCLGVHIEFGFKWENFYCE 120
Db 61 SEKHVWIGLVRVQGVKQCNSEWSDGSSSVHYDNLQENKRCYCLEKRAFFTWNVYCG 120

QY 121 QDDPFVCE 128
Db 121 HEYFVCK 128

RESULT 3
BOTA BOTJA STANDARD; PRT; 133 AA.
ID P22029;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Botrocetin, alpha chain (Platelet coagglutinin).
OS Bothrops jararaca (Jararaca).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidodactylus; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Bothrops.
OX NCBI_TaxID=8724;
RN [1]
RP SEQUENCE, AND DISULFIDE BONDS.
RC TISSUE=Venom;
RX MEDLINE=93157385; PubMed=8430107;
RA Usami Y., Fujimura Y., Suzuki M., Ozeki Y., Nishio K., Fukui H.,
RA Titani K.;
RT "Primary structure of two-chain botrocetin, a von Willebrand factor modulator purified from the venom of Bothrops jararaca.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:928-932(1993).
RN [2]
RP SEQUENCE OF 1-40.
RC TISSUE=Venom;
RX MEDLINE=9129280; PubMed=1993206;
RA Fujimura Y., Titani K., Usami Y., Suzuki M., Oyama R., Matsui T.,
RA Fukui H., Sugimoto M., Ruggeri Z.M.;
RT "Isolation and chemical characterization of two structurally and functionally distinct forms of botrocetin, the platelet coagglutinin isolated from the venom of Bothrops jararaca.";
RL Biochemistry 30:1957-1964(1991).
RN [3]

```

RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
RX MEDLINE=22118144; PubMed=12121649;
RA Fukuda K., Doggett T.A., Bankston L.A., Cruz M.A., Diacovo T.G.,
RA Liddington R.C.;
RT "Structural basis of von Willebrand factor activation by the snake
toxin botrocetin.";
RL STRUCTURE 10:943-950 (2002).
CC -!- FUNCTION: Two-chain Botrocetin forms an activated complex with
CC vWF, and the complex then binds to platelet GPIb, resulting in
CC platelet agglutination.
CC -!- FUNCTION: There are two distinct forms of the von Willebrand
CC factor-dependent platelet coagglutinin. The dimeric form is
CC 34-times more active than the one-chain Botrocetin in promoting
CC vWF binding to platelets.
CC -!- SUBUNIT: Disulfide-linked dimer of an alpha and a beta chain.
CC vWF and Botrocetin form a soluble complex.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC PIR: A47267; A47267.
DR PDB: 1IJK; 17-JUL-02.
DR PDB: 1FVU; 14-FEB-01.
DR InterPro; IPR001304; AntifreezeII.
DR InterPro; IPR001304; Lectin C.
DR Pfam; PF00059; lectin c; 1.
DR PRINTS; PR00356; ANTIFREEZEII.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW Lectin; 3D-structure.
FT DISULFID 2 13
FT DISULFID 30 128
FT DISULFID 80 80
FT DISULFID 103 120
SQ SEQUENCE 133 AA; 15215 MW; E4CF4502946AC74B CRC64;

Query Match 59.1%; Score 422.5; DB 1; Length 133;
Best Local Similarity 56.9%; Pred. No. 2.3e-33;
Matches 74; Conservative 20; Mismatches 35; Indels 1; Gaps 1;

QY 1 DCSDDWSYEGHCYKVKQSKTWTDAESFCTKQVNGHLSIESGSEADFGVQLIAQKI 59
Db 1 DCPGMSAYDQCYRVFNPQWDAERFCAQASGHLVSIETGEADFVQLISENQTQS 60
QY 60 KSAKHVWIGLRAQNKQCSIESGSSISKENWIEESKCLGVHETGFHWENFYCEQ 119
Db 61 QSSDLVWIGLRAQNKQCSIESGSSISKENWIEESKCLGVHETGFHWENFYCEQ 120
QY 120 EQDPFVCEA 129
Db 121 AQKNPFVCKS 130

RESULT 4
ID AB2A TRIAB STANDARD; PRT; 134 AA.
AC P81112;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Alboaggregin A subunit 2.
OS Trimeresurus albolabris (White-lipped pit viper).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Trimeresurus.
OX NCBI_TaxID=8765;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=98189535; PubMed=9531050;
RA Kowalska M.A., Tan L., Holt J.C., Peng M., Karczewski J.,
RA Calvete J.J., Niewiarowski S.;
RT "Alboaggregin A and B. Structure and interaction with human
platelets.";
RL Trimeresurus albolabris (White-lipped pit viper).
CC -!- FUNCTION: Heterodimer of alpha and beta subunits, disulfide-linked.
CC agglutination.
CC -!- SUBUNIT: Heterodimer of alpha and beta subunits, disulfide-linked.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC HSSP; P23806; IIXX.
DR InterPro; IPR001304; Lectin C.
DR Pfam; PF00059; lectin c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW platelets.;

RL Thromb. Haemost. 79:609-613 (1998).
CC -!- FUNCTION: Binds to platelet GPIb/IX receptor system and stimulates
CC agglutination.
CC -!- SUBUNIT: Heterotetramer of the subunits 1, 2, 3 and 4,
CC disulfide-linked.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
DR HSSP; P23806; IIXX.
DR InterPro; IPR001304; Lectin C.
DR InterPro; IPR003990; Pancreatins_ac.
DR Pfam; PF00059; lectin c; 1.
DR PRINTS; PR01504; PNCREATINSAP.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW Lectin.
FT DOMAIN 3 131 C-TYPE LECTIN (LONG FORM).
FT DISULFID 4 15 BY SIMILARITY.
FT DISULFID 32 129 BY SIMILARITY.
FT DISULFID 104 121 BY SIMILARITY.
SQ SEQUENCE 134 AA; 15556 MW; 5F9D71FC86DE2435 CRC64;

Query Match 57.1%; Score 408; DB 1; Length 134;
Best Local Similarity 56.7%; Pred. No. 5.4e-32;
Matches 72; Conservative 21; Mismatches 34; Indels 0; Gaps 0;

QY 2 CSDWSYEGHCYKVKQSKTWTDAESFCTKQVNGHLSIESGSEADFGVQLIAQTKS 61
Db 4 CLPGMSAYDQCYRVFNPQWDAERFCAQASGHLVSIETGEADFVQLISENQTQS 63
QY 62 AKIHVWIGLRAQNKQCSIESGSSISKENWIEESKCLGVHETGFHWENFYCEQ 121
Db 64 EKHVWIGLRAQNKQCSIESGSSISKENWIEESKCLGVHETGFHWENFYCEQ 123
QY 122 QDPFVCE 128
Db 124 LNPVCK 130

RESULT 5
ID AB2A TRIAB STANDARD; PRT; 132 AA.
AC P81115;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Alboaggregin B alpha subunit.
OS Trimeresurus albolabris (White-lipped pit viper).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Trimeresurus.
OX NCBI_TaxID=8765;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=98189535; PubMed=9531050;
RA Kowalska M.A., Tan L., Holt J.C., Peng M., Karczewski J.,
RA Calvete J.J., Niewiarowski S.;
RT "Alboaggregin A and B. Structure and interaction with human
platelets.";
RL Trimeresurus albolabris (White-lipped pit viper).
CC -!- FUNCTION: Binds to platelet GPIb/IX receptor system and stimulates
CC agglutination.
CC -!- SUBUNIT: Heterodimer of alpha and beta subunits, disulfide-linked.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC HSSP; P23806; IIXX.
DR InterPro; IPR001304; Lectin C.
DR Pfam; PF00059; lectin c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW Lectin.

FT DOMAIN 1 129 C-TYPE LECTIN (LONG FORM).
 FT DISULFID 2 13 BY SIMILARITY.
 FT DISULFID 30 127 BY SIMILARITY.
 FT DISULFID 79 79 INTERCHAIN (WITH C-75 IN BETA CHAIN)
 (BY SIMILARITY).
 FT DISULFID 102 119 BY SIMILARITY.
 FT DISULFID 132 AA; 15419 MW; 6FAE64820383F16F CRC64;
 SQ SEQUENCE 54.3%; Score 388; DB 1; Length 132;
 Query Match 52.3%; Pred. No. 4.2e-30;
 Best Local Similarity 21; Mismatches 40; Indels 0; Gaps 0;
 Matches 67; Conservative 21; Mismatches 40; Indels 0; Gaps 0;
 Qy 1 DCSDMSVEGHCYKVFQSKTWTDAESFCTKVNGGHLVSTESSGEADFVGLIAQIK 60
 Db 1 DCPDMSVFKYQVIVKELTWEDAEPCQANDGHLVSTESREAVFAELLSENVX 60
 Qy 61 SAKHVWIGLRAQNKQCSIEWSDGSSISKENWIEESKCLGVHIETGFHKWENFYCE 120
 Db 61 XXXVHWIGLVSQNKXQCSSEWSGSSVTYENLXXXXKCFVLKESFRTWSNVYCE 120
 Qy 121 QDPFVCE 128
 Db 121 QKHPMCK 128

RESULT 6

CVXA CRODU STANDARD; PRT; 158 AA.
 ID ECHA ECHCA AC 093426;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Convulxin alpha precursor (CVX alpha).
 OS Crotalus durissus terrificus (South American rattlesnake).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
 OC Viperidae; Crotalinae; Crotalus.
 OX NCBI_TaxID=8732;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 24-48; 61-68; 108-120; 139-145 AND 149-153.
 RC TISSUE=Venom gland;
 RX MEDLINE=98324901; PubMed=9657980;
 RA Leduc M., Bon C.;
 RT "Cloning of subunits of convulxin, a collagen-like platelet-aggregating protein from Crotalus durissus terrificus venom.";
 RL Biochem. J. 333:389-393(1998).
 CC -!- FUNCTION: Binds to the platelet and collagen receptor, glycoprotein VI (GPVI).
 CC -!- SUBUNIT: Heterodimer of three alpha chains and three beta chains; disulfide-linked.
 CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Y16348; CAA76181.1; --
 DR HSP; P23806; I1XX.
 DR InterPro; IPR001304; Lectin.C.
 DR InterPro; IPR003990; Pancreatis_ac.
 DR Pfam; PF00059; lectin c; 1.
 DR PRINTS; PR01504; PNCREATISAP.
 DR SMART; SM00034; CLECT.1.
 DR PROSITE; PS00615; C-TYPE LECTIN 1; FALSE_NEG.
 DR PROSITE; PS50041; C-TYPE LECTIN 2; 1.
 KW Lectin; Glycoprotein; Signal.
 FT SIGNAL 1 23
 FT CHAIN 24 158 CONVULXIN ALPHA.
 FT

FT DOMAIN 34 153 C-TYPE LECTIN.
 FT DISULFID 27 38 BY SIMILARITY.
 FT DISULFID 55 152 BY SIMILARITY.
 FT DISULFID 104 104 INTERCHAIN (WITH C-100 IN BETA CHAIN)
 (POTENTIAL).
 FT DISULFID 127 144 BY SIMILARITY.
 FT DISULFID 158 158 INTERCHAIN (WITH C-26 IN BETA CHAIN)
 (POTENTIAL).
 SQ SEQUENCE 158 AA; 18141 MW; 949F9C6D673E2318 CRC64;
 Query Match 50.9%; Score 364; DB 1; Length 158;
 Best Local Similarity 52.8%; Pred. No. 9.6e-28;
 Matches 67; Conservative 18; Mismatches 42; Indels 0; Gaps 0;
 Qy 2 CSSDMSVEGHCYKVFQSKTWTDAESFCTKVNGGHLVSTESSGEADFVGLIAQIKS 61
 Db 27 CPDMSVYDQHCYKIFNEEMWEDAEFCTKQAGHLVSIKSAEADFVAMVTONIEE 86
 Qy 62 AKHVWIGLRAQNKQCSIEWSDGSSISKENWIEESKCLGVHIETGFHKWENFYCEQ 121
 Db 87 SFSVHSIGLRVQNKQCSIEWSDGSSVSYDNLDDLYITKCSLLKKTGFRKFWVASCIG 146
 Qy 122 QDPFVCE 128
 Db 147 KIPFVCK 153

RESULT 7

ECHA ECHCA STANDARD; PRT; 133 AA.
 ID ECHA ECHCA AC P81017;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Echicetin alpha subunit.
 OS Echinocarinatus (Saw-scaled viper).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
 OC Viperidae; Viperinae; Echis.
 OX NCBI_TaxID=40353;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Venom;
 RX MEDLINE=97250657; PubMed=9163349;
 RA Polgar J., Magnenat E.M., Peitsch M.C., Wells T.N.C., Saqi M.S.A., Clenetsen K.J.;
 RT "Amino acid sequence of the alpha subunit and computer modelling of the alpha and beta subunits of echicetin from the venom of Echis carinatus (saw-scaled viper).";
 RL Biochem. J. 323:533-537(1997).
 RN [2]
 RP CHARACTERIZATION.
 RC TISSUE=Venom;
 RX MEDLINE=9324424; PubMed=8481512;
 RA Peng M., Lu W., Bevilgia L., Niewiarowski S., Kirby E.P.;
 RT "Echicetin: a snake venom protein that inhibits binding of von Willebrand factor and also aggregates to platelet glycoprotein Ib.";
 RL Blood 81:2321-2328(1993).
 CC -!- FUNCTION: Binds to platelet GPIb and inhibits platelet agglutination.
 CC -!- SUBUNIT: Heterodimer of alpha and beta subunits.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
 DR HSP; P23806; I1XX.
 DR InterPro; IPR001304; Lectin.C.
 DR InterPro; IPR003990; Pancreatis_ac.
 DR Pfam; PF00059; lectin c; 1.
 DR PRINTS; PR01504; PNCREATISAP.
 DR SMART; SM00034; CLECT.1.
 DR PROSITE; PS00615; C-TYPE LECTIN 1; FALSE_NEG.
 DR PROSITE; PS50041; C-TYPE LECTIN 2; 1.
 KW Lectin.
 FT SIGNAL 4 15 BY SIMILARITY.
 FT DISULFID

ID	IXB_TITLE	STANDARD;	PRT;	146 AA.
DT	P23807; Q91247;			
DT	01-NOV-1991 (Rel. 20, Created)			
DT	15-JUL-1998 (Rel. 36, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Coagulation factor IX/factor X-binding protein B chain precursor (IX/X-BP).			
OS	Trimeresurus flavoviridis (Habu).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;			
OC	Viperidae; Crotalinae; Trimeresurus.			
OX	NCBI_TaxID=88087;			
RN	[1]			
RN	SEQUENCE FROM N.A.			
RP	MEDLINE=96184662; PubMed=8645314;			
RP	MATSUZAKI R., Yoshihara E., Yamada M., Shima K., Atoda H., Morita T.;			
RP	"CDNA cloning of IX/X-BP, a heterogeneous two-chain anticoagulant protein from snake venom.";			
RP	Biochem. Biophys. Res. Commun. 220:382-387(1996).			
RN	[2]			
RP	SEQUENCE OF 24-146.			
RP	TISSUE=Venom;			
RC	MEDLINE=91332000; PubMed=1831197;			
RA	Atoda H., Hyuga M., Morita T.;			
RA	"The primary structure of coagulation factor IX/factor X-binding protein isolated from the venom of Trimeresurus flavoviridis.			
RT	Homology with asialoglycoprotein receptors, proteoglycan core protein, tetranectin, and lymphocyte Fc epsilon receptor for immunoglobulin E.";			
RT	J. Biol. Chem. 266:14903-14911(1991).			
RN	[3]			
RP	X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).			
RP	MEDLINE=97331317; PubMed=9187649;			
RA	Mizuno H., Fujimoto Z., Koizumi M., Kano H., Atoda H., Morita T.;			
RT	"Structure of coagulation factors IX/X-binding protein, a heterodimer of C-type lectin domains.";			
RT	Nat. Struct. Biol. 4:438-441(1997).			
CC	1- FUNCTION: ANTICOAGULANT PROTEIN WHICH BINDS WITH FACTOR IX AND FACTOR X IN THE PRESENCE OF CALCIUM WITH A 1 TO 1 STOICHIOMETRY.			
CC	1- SUBUNIT: HETERODIMER OF CHAINS A AND B LINKED BY A DISULFIDE BOND.			
CC	1- SUBCELLULAR LOCATION: Secreted.			
CC	1- MISCELLANEOUS: CALCIUM IS REQUIRED FOR LIGAND BINDING.			
CC	1- SIMILARITY: Contains 1 C-type lectin family domain.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch).			
DR	EMBL; D83332; BAA11888.1; ..			
DR	PIR; JC4691; JC4691			
DR	PDB; 1LXX; 06-MAY-98.			
DR	PDB; 1B33; 16-AUG-99.			
DR	InterPro; IPR001304; Lectin_C.			
DR	Pfam; PF00059; Lectin_c; 1.			
DR	SMART; SM00034; CLECT; 1.			
DR	PROSITE; PS00615; C_TYPE_LECTIN_1; 1.			
DR	PROSITE; PS00041; C_TYPE_LECTIN_2; 1.			
KW	Lectin; Calcium; Signal; 3D-structure.			
FT	SIGNAL			
FT	1 23			
FT	CHAIN 24 146			
FT	DOMAIN 24 144			
FT	DISULFID 25 36			
FT	DISULFID 53 142			
FT	DISULFID 98 98			
FT	DISULFID 119 134			
FT	TURN 27 28			
FT	TURN 30 32			
FT	TURN 33 34			
FT	TURN 33 34			
FT	COAGULATION FACTOR IX/FACTOR X-BINDING PROTEIN B CHAIN.			
FT	C-TYPE LECTIN (LONG FORM).			
FT	BY SIMILARITY.			
FT	BY SIMILARITY.			
FT	INTERCHAIN (WITH C-102 OF A CHAIN).			
FT	BY SIMILARITY.			

```

FT DISULFID 95 112 BY SIMILARITY.
SQ SEQUENCE 127 AA; 15162 MW; B5DA100D383E3547 CRC64;

Query Match
Best Local Similarity 38.3%; Score 273.5; DB 1; Length 127;
Matches 53; Conservative 21; Mismatches 45; Indels 9; Gaps 2;

QY 1 DCSWSSYEGHCYKVFQKSKTWTDAESFCTKVNGHLSVSISSGADPFGQLIAQIK 60
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 3 ECFSGSSYDRICYKPFKQEMTWADAEFCQAKGRHLSVETALEASFVNVIYANKE 62
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 61 SAKIHWIGLRAQNKQKCSIEWSGSSISKNNWIEESKCLGVHIETGFHKWENFYCE 120
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 63 YLTRYIWIGLRVQNRQGPC-----SSYSENVLDPDP--CFWYSDRLREWFKYDCE 113
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 121 QDPFVCE 128
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 114 QHSPICK 121
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 12
RHCX AGKRH
ID RHCX AGKRH STANDARD; PRT; 129 AA.
AC P81398;
DT 15-JUL-1998 (Rel. 36, Created)
DT 28-FEB-2003 (Rel. 36, Last sequence update)
DE Rhodocetin beta subunit.
OS Agkistrodon rhodostoma (Malayan pit viper) (Calloselasma rhodostoma).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Calloselasma.
OX NCBI_TaxID=8717;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 24-53; 99-109 AND 126-145.
RC TISSUE=Venom;
RX MEDLINE=99303998; PubMed=10360956;
RA Wang R., Kini R.M., Chung M.C.M.;
RT "Rhodocetin, a novel platelet aggregation inhibitor from the venom of
RT Calloselasma rhodostoma (Malayan pit viper): synergistic and
RT noncovalent interaction between its subunits.";
RL Biochemistry 38:7584-7593(1999).
CC -!- FUNCTION: A potent inhibitor of collagen-induced platelet
CC aggregation. Individually, neither subunit inhibits platelet
CC aggregation. Both subunits are essential.
CC -!- SUBUNIT: Heterodimer of one alpha and one beta subunit held
CC together by noncovalent interactions rather than by intersubunit
CC disulfide bridges.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- MASS SPECTROMETRY: MW=15184.53; MW ERR=2.74; METHOD=Electrospray.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
DR HSSP; P23806; IIXX.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; FALSE_NEG.
DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
KW Lectin.
FT DOMAIN 3 125 C-TYPE LECTIN (LONG FORM).
FT DISULFID 4 15 BY SIMILARITY.
FT DISULFID 98 115 BY SIMILARITY.
SQ SEQUENCE 129 AA; 15190 MW; BF4B318FAAC807AE CRC64;

Query Match
Best Local Similarity 37.6%; Score 269; DB 1; Length 129;
Matches 50; Conservative 23; Mismatches 48; Indels 6; Gaps 1;

QY 2 CSDSWSSYEGHCYKVFQKSKTWTDAESFCTKVNGHLSVSISSGADPFGQLIAQIKS 61
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 4 CPTWSASKLYCYKPFKEKKTWAEARFCQAKQENGLHLSVIGSAEADFLLDVLIVYDFK 63
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 62 AKIHWIGLRAQNKQKCSIEWSGSSISKNNWIEESKCLGVHIETGFHKWENFYCEQ 121
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

```

```

Db 64 QRYRAWTGLTERN-----LKTNGASVSVENLYEPYIRKCFVVPQWEGSKWKYKADCE 117
QY 122 QDPFVCE 128
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 118 KNAFLCK 124
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 13
CVXB_CRODU
ID CVXB_CRODU STANDARD; PRT; 148 AA.
AC Q93427;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Convulxin beta precursor (CVX beta).
OS Crotalus durissus terrificus (South American rattlesnake).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Crotalus.
OX NCBI_TaxID=8732;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 24-53; 99-109 AND 126-145.
RC TISSUE=Venom gland;
RX MEDLINE=98324901; PubMed=9657980;
RA Leduc M., Bon C.;
RT "Cloning of subunits of convulxin, a collagen-like platelet-
RT aggregating protein from Crotalus durissus terrificus venom.";
RL Biochem. J. 333:389-393(1998).
CC -!- FUNCTION: Binds to the platelet and collagen receptor,
CC glycoprotein VI (GPVI).
CC -!- SUBUNIT: Heterohexamer of three alpha chains and three beta
CC chains; disulfide-linked.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; Y16349; CAA76182.1; -.
DR HSSP; P23807; IIXX.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
KW Lectin; Glycoprotein; Signal.
FT SIGNAL 1 23
FT CHAIN 24 148 CONVULXIN BETA.
FT DOMAIN 34 145 C-TYPE LECTIN.
FT DISULFID 26 26 INTERCHAIN (WITH C-158 IN ALPHA CHAIN)
(POTENTIAL).
FT DISULFID 27 38 BY SIMILARITY.
FT DISULFID 55 144 BY SIMILARITY.
FT DISULFID 100 100 INTERCHAIN (WITH C-104 IN ALPHA CHAIN)
(POTENTIAL).
FT DISULFID 121 136 BY SIMILARITY.
SQ SEQUENCE 148 AA; 17402 MW; 94D7E3E1BC693B9F CRC64;

Query Match
Best Local Similarity 36.9%; Score 264; DB 1; Length 148;
Matches 57; Conservative 17; Mismatches 44; Indels 10; Gaps 5;

QY 2 CSDSWSSYEGHCYKVFQKSKTWTDAESFCTKVNGHLSVSISSGADPFGQLIAQIKS 61
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 27 CFSWSSYDRICYKVFQKQEMTWADAEKFCCTQHTSHLSVSISSGADPFGQLIAQIKS 86
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 62 AKIHWIGLRAQNKQKCSIEWSGSSISKNNWIEESKCLGVHIETGF-HKWNFYCE 120
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 87 T--FFWIG--ANNIWNKNQWSDGTPKEKWEH--PECL--ISRFTDNQWLSAPCS 137
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

```

Best Local Similarity 42.2%; Pred. No. 5.4e-18;
Matches 54; Conservative 17; Mismatches 51; Indels 6; Gaps 3;
QY 1 DCSSDSSSYEGHCYKVFQSKTWTDAESFCTQVNGHGLVSISSGSEADPVGOLIAQKIK 60
Db 1 DCPDSSSYEGHCYKVFQSKTWTDAESFCTQVNGHGLVSISSGSEADPVGOLIAQKIK 60
QY 61 SAKIHWIGLRAQNKKEKQCSIEWSDGSSISKENWIEESKCLGVHITGFHKWENFYC 120
Db 61 GDV--VWIGL--SDVWVKCFEFTDGMEDFYDDYLLAEYEC--VASKPTNNKWWIIPCT 114
QY 121 QQDPFFVCE 128
Db 115 RFXNFVCE 122
RESULT 15
CHBB CROHO
ID CHBB CROHO STANDARD; PRT; 117 AA.
AC P81509;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE CH-B beta subunit.
OS Crotalus horridus horridus (timber rattlesnake).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Crotalus.
OX NCBI_TaxID=8747;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
EX MEDLINE=96420502; PubMed=8823201;
RA Andrews R.K., Kroll M.H., Ward C.M., Rose J.W., Scarborough R.M.,
RA Smith A.I., Lopez J.A., Berndt M.C.;
RT "Binding of a novel 50-kilodalton albosagregin from Trimeresurus
RT albolabris and related viper venom proteins to the platelet membrane
RT glycoprotein Ib-IX-V complex. Effect on platelet aggregation and
RT glycoprotein Ib-mediated platelet activation.";
RL Biochemistry 35:12629-12639(1996).
CC -!- FUNCTION: Binds to platelet GPIb/IX receptor system, inhibits VWF
CC binding, and stimulates agglutination.
CC -!- SUBUNIT: Heterodimer of alpha and beta subunits; disulfide-linked.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
DR SMART: SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
DR PROSITE; PS00041; C_TYPE_LLECTIN_2; 1.
KW Lectin.
FT DOMAIN 9 116 C-TYPE LECTIN.
FT DISULFID 2 13 BY SIMILARITY.
FT DISULFID 30 115 BY SIMILARITY.
FT DISULFID 92 92 INTERCHAIN (WITH C-81 IN ALPHA CHAIN)
FT (POTENTIAL).
FT DISULFID 95 107 BY SIMILARITY.
SQ SEQUENCE 117 AA; 13888 MW; 07835BEC61E9EAD CRC64;
Query Match 36.2%; Score 259; DB 1; Length 117;
Best Local Similarity 40.3%; Pred. No. 6.3e-18;
Matches 52; Conservative 21; Mismatches 42; Indels 14; Gaps 4;
QY 1 DCSSDSSSYEGHCYKVFQSKTWTDAESFCTQVNGHGLVSISSGSEADPVGOLIAQKIK 60
Db 1 DCPDSSSYEGHCYKVFQSKTWTDAESFCTQVNGHGLVSISSGSEADPVGOLIAQKIK 59
QY 61 SAKIHWIGLRAQNKKEKQCSIEWSDGSSISKENWIEESKCLGVHITGFHKWENFYC 119
Db 60 -----FWMGWRDITWNER--LQWSDGTKNVYKAWSAEPECIVCRATD-----NQWLSTSC 107
QY 120 EQQDPFFVCE 128
Db 108 SKTHNVVCK 116

Best Local Similarity 42.2%; Pred. No. 5.4e-18;
Matches 54; Conservative 17; Mismatches 51; Indels 6; Gaps 3;
QY 1 DCSSDSSSYEGHCYKVFQSKTWTDAESFCTQVNGHGLVSISSGSEADPVGOLIAQKIK 60
Db 1 DCPDSSSYEGHCYKVFQSKTWTDAESFCTQVNGHGLVSISSGSEADPVGOLIAQKIK 60
QY 61 SAKIHWIGLRAQNKKEKQCSIEWSDGSSISKENWIEESKCLGVHITGFHKWENFYC 120
Db 61 GDV--VWIGL--SDVWVKCFEFTDGMEDFYDDYLLAEYEC--VASKPTNNKWWIIPCT 114
QY 121 QQDPFFVCE 128
Db 115 RFXNFVCE 122
RESULT 15
CHBB CROHO
ID CHBB CROHO STANDARD; PRT; 125 AA.
AC P22030;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Botrocetin, beta chain (Platelet coagglutinin).
OS Bothrops jararaca (Jararaca).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Bothrops.
OX NCBI_TaxID=8724;
RN [1]
RP SEQUENCE, AND DISULFIDE BONDS.
RC TISSUE=Venom;
EX MEDLINE=93157385; PubMed=8430107;
RA Usami Y., Fujimura Y., Suzuki M., Ozeki Y., Nishio K., Fukui H.,
RA Titani K.;
RT "Primary structure of two-chain botrocetin, a von Willebrand factor
RT modulator purified from the venom of Bothrops jararaca.";
RL proc. Natl. Acad. Sci. U.S.A. 90:928-932(1993).
RN [2]
RP SEQUENCE OF 1-40.
RC TISSUE=Venom;
EX MEDLINE=91129280; PubMed=1993206;
RA Fujimura Y., Titani K., Usami Y., Suzuki M., Oyama R., Matsui T.,
RA Fukui H., Sugimoto M., Ruggeri Z.M.;
RT "Isolation and chemical characterization of two structurally and
RT functionally distinct forms of botrocetin, the platelet coagglutinin
RT isolated from the venom of Bothrops jararaca.";
RL Biochemistry 30:1957-1964(1991).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
EX MEDLINE=22118144; PubMed=12121649;
RA Fukuda K., Doggett T.A., Bankston L.A., Cruz M.A., Diacovo T.G.,
RA Liddington R.C.;
RT "Structural basis of von Willebrand factor activation by the snake
RT toxin botrocetin.";
RL Structure 10:943-950(2002).
CC -!- FUNCTION: Two-chain Botrocetin forms an activated complex with
CC VWF, and the complex then binds to platelet GPIb, resulting in
CC platelet agglutination.
CC -!- FUNCTION: There are two distinct forms of the von Willebrand
CC factor-dependent platelet coagglutinin. The dimeric form is
CC 34-times more active than the one-chain Botrocetin in promoting
CC VWF binding to platelets.
CC -!- SUBUNIT: Disulfide-linked dimer of an alpha and a beta chain.
CC VWF and Botrocetin form a soluble complex.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
DR PIR; B47267; B47267.
DR PDB; 1JUK; 17-JUL-02.
DR PDB; 1FVU; 14-FEB-01.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; FALSE_NEG.
DR PROSITE; PS00041; C_TYPE_LLECTIN_2; 1.
KW Lectin; 3D-structure.
FT DISULFID 2 13
FT DISULFID 30 121 INTERCHAIN (WITH C-80 IN ALPHA CHAIN).
FT DISULFID 75 75
FT DISULFID 98 113
SQ SEQUENCE 125 AA; 15037 MW; 1ED2027ED817FCA0 CRC64;
Query Match 36.4%; Score 260; DB 1; Length 125;

Tue Dec 9' 09:26:55 2003

us-09-938-114-2.rsp

Page 9

Search completed: December 8, 2003, 09:53:09
Job time : 17.2876 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:49:24 ; Search time 62.6571 Seconds
(without alignments)
531.285 Million cell updates/sec

Title: US-09-938-114-2

Perfect score: 715

Sequence: 1 DCSSDSSVSEGHCHYKVFQKS.....GFHKWENFYCEQDDPFVCEA 129

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_virus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	701	98.0	152	13	Q9IAM1	Q9iam1 agkistrodon
2	657	91.9	152	13	Q9DEF9	Q9def9 agkistrodon
3	657	91.9	152	13	Q8JTW0	Q8jtw0 agkistrodon
4	496	69.4	146	13	Q98UJ0	Q98uj0 trimeresuru
5	475	66.4	152	13	Q9DGJ9	Q9dgj9 agkistrodon
6	460	64.3	154	13	Q9YGN5	Q9ygn5 agkistrodon
7	439	61.4	158	13	Q8UVC6	Q8uvc6 agkistrodon
8	439	61.4	158	13	Q8AYAS	Q8ayas agkistrodon
9	420.5	58.8	136	13	Q9I841	Q9i841 agkistrodon
10	394	55.1	158	13	Q8JGT7	Q8jgt7 trimeresuru
11	391	54.7	158	13	Q8AV97	Q8av97 trimeresuru
12	390	54.5	131	13	Q9PSM9	Q9psm9 echis carin
13	382.5	53.5	142	13	Q9PSM6	Q9psm6 bothrops ja
14	380.5	53.2	155	13	Q8JTW6	Q8jtw6 agkistrodon
15	377.5	52.8	155	13	Q9DEA2	Q9dea2 agkistrodon
16	375	52.4	154	13	Q8JTV9	Q8jtv9 agkistrodon

17	355.5	49.7	157	13	Q9YGG9	Q9ygg9 agkistrodon
18	290	40.6	146	13	Q9DG31	Q9dg31 agkistrodon
19	290	40.6	146	13	Q8AYA4	Q8aya4 agkistrodon
20	289	40.4	146	13	Q8JTV7	Q8jtv7 agkistrodon
21	282	39.4	124	13	Q98SM5	Q98sm5 agkistrodon
22	280	39.2	146	13	Q9I840	Q9i840 agkistrodon
23	278	38.9	146	13	Q9IAM0	Q9iam0 agkistrodon
24	277.5	38.8	125	13	Q9PSM8	Q9psm8 echis carin
25	277	38.7	146	13	Q9DEF8	Q9def8 agkistrodon
26	274	38.3	146	13	Q8JTW1	Q8jtw1 agkistrodon
27	273	38.2	146	13	Q9YI92	Q9yi92 agkistrodon
28	272.5	38.1	145	13	Q9YGN4	Q9ygn4 agkistrodon
29	271.5	38.0	149	13	Q8UVC7	Q8uvc7 agkistrodon
30	268.5	37.6	149	13	Q8AYA3	Q8aya3 agkistrodon
31	256	35.8	123	13	Q9PSM5	Q9psm5 bothrops ja
32	256	35.8	155	13	Q8JTV8	Q8jtv8 agkistrodon
33	247	34.5	146	13	Q9DEA1	Q9dea1 agkistrodon
34	246	34.4	146	13	Q8JGT6	Q8jgt6 trimeresuru
35	239	33.4	124	13	Q90WL9	Q90wl9 agkistrodon
36	232	32.4	148	13	Q8AV98	Q8av98 trimeresuru
37	227.5	31.8	151	13	Q8JTW2	Q8jtw2 agkistrodon
38	215.5	30.1	158	13	Q90WL7	Q90wl7 bungarus fa
39	214.5	30.0	135	13	Q9PSM4	Q9psm4 lachesis mu
40	206.5	28.9	158	13	Q90WI6	Q90wi6 bungarus mu
41	205.5	28.7	158	13	Q90WI8	Q90wi8 bungarus fa
42	204.5	28.6	135	13	Q9PSN0	Q9psn0 bitis ariet
43	201.5	28.2	1479	4	Q9YSP9	Q9y5p9 homo sapien
44	201.5	28.2	1479	4	Q9UBG0	Q9ubg0 homo sapien
45	199.5	27.9	719	6	Q82623	Q82623 bos taurus

ALIGNMENTS

RESULT 1

Q9IAM1
ID Q9IAM1 PRELIMINARY; PRT; 152 AA.
AC Q9IAM1;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Agkisutacin A chain.
OS Agkisutacin acutus (Hundred-pace snake) (Deinagkistrodon acutus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;
OC Viperidae; Crotalinae; Deinagkistrodon.
OX NCBI_TaxID=36307;
[1]
RN
RP SEQUENCE FROM N.A.
RC TISSUE=Venom;
RA Yu H.-X., Xiang K.-J., Liu J.;
RT "cDNA sequencing and analysis of eleven C-type lectin-like protein subunits from Agkistrodon acutus";
RL Chin. J. Biochem. Biophys. 0:0-0(2002).
DR EMBL; AF176420; AAF26286.2; -;
DR HSSP; P23806; 11XX.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF000059; lectin_c; 1.
DR SMART; SM00034; CLECT_1.
DR PROSITE; PS000615; C_TYPE_LLECTIN_1; 1.
DR PROSITE; PS00041; C_TYPE_LLECTIN_2; 1.
SQ SEQUENCE 152 AA; 17109 MW; 76A0F636DBFD7AB CRC64;

Query Match 98.0%; Score 701; DB 13; Length 152;
Best Local Similarity 98.4%; Pred. No. 2.6e-66;
Matches 127; Conservativity 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 DCSSDSSVSEGHCHYKVFQKSWTDAESFCTKQVNGHLVSISSGEADPVGQLIAQKIK 60
Db 24 DCSSGSSVSEGHCHYKVFQKSWTDAESFCTKQVNGHLVSISSGEADPVGQLIAQKIK 83
QY 61 SAKHWIGLRANKEKQCSIEWSDGSSISKENWIEESKCLGVHETGFHKWENFYCE 120
|||||

RT "A chain of ACF 1/2 from Deinagkistrodon acutus.";

```

QY 121 QDDPFVCEA 129
Db 138 QONPFVCEA 146

RESULT 5
ID Q9DG39 PRELIMINARY; PRT; 152 AA.
AC Q9DG39;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Halyxin A-chain precursor.
GN HXNA.
OS Agkistrodon halys pallas (Chinese water moccasin) (Gloydus halys pallas).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroides;
OC Viperidae; Crotalinae; Gloydus.
OX NCBI_TaxID=8714;
RN [1]
RP SEQUENCE FROM N.A.
RA Koo B.H., Sohn Y.D., Kim D.S., Jang Y.S., Chung K.H.;
RT "A novel coagulation factor Xa inhibitor from Korean snake
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF190827; AAG17178.1; -.
DR HSSP; P23806; 11XX.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR PRINTS; PR01504; PNCREATITSAP.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
RW Signal.
FT SIGNAL
FT CHAIN
FT CHAIN
FT SEQUENCE
FT SEQUENCE 152 AA; 17455 MW; BDD74D1DC280C28D CRC64;

Query Match 66.4%; Score 475; DB 13; Length 152;
Best Local Similarity 63.6%; Pred. No. 2.1e-42;
Matches 82; Conservative 17; Mismatches 30; Indels 0; Gaps 0;

QY 1 DCSSDSSSYEGHCYKVKQSKTWTDAESFCTKVQNGHLSIESSEADPVGOLIAQKIK 60
Db 24 DCPSGSSSYEGHCYNIPLFKTWAERFCRKQVGAHLVSISSSEADPVAQLVSENMK 83

QY 61 SAKIHVWIGLRAQNKQCSIEWSGSSISKENWIEESKCLGVHIEGFKHWFNFCYCE 120
Db 84 RYGIYIWIGLVRGKKQCSQSSQSSVYQNWIEAESTCLGLQKETEFKRWFIYCG 143

QY 121 QDDPFVCEA 129
Db 144 ERNPFVCEA 152

RESULT 6
Q9YGN5 PRELIMINARY; PRT; 154 AA.
ID Q9YGN5
AC Q9YGN5;
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Fibrinogen clotting inhibitor A chain.
OS Agkistrodon halys breviceaudus (Korean slamosa snake) (Gloydus halys breviceaudus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroides;
OX Viperidae; Crotalinae; Gloydus.
OX NCBI_TaxID=66175;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RA Kim D.S., Koh Y.S.;
RT "Purification and molecular cloning of snake venom fibrin clotting
RL inhibitor.";
DR EMBL; AF125309; AAD18055.1; -.
DR HSSP; P23806; 11XX.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
SQ SEQUENCE 154 AA; 17293 MW; 8D06F7DDFA0D140D CRC64;

Query Match 64.3%; Score 460; DB 13; Length 154;
Best Local Similarity 60.2%; Pred. No. 8.1e-41;
Matches 77; Conservative 20; Mismatches 31; Indels 0; Gaps 0;

QY 2 CSSDSSSYEGHCYKVKQSKTWTDAESFCTKVQNGHLSIESSEADPVGOLIAQKIKS 61
Db 27 CPSCGSSNGHCYQAFNQMRTWEDAEERFCSAQAGKGLVSIETRAEADFVAHVVAERLET 86

QY 62 AKIHVWIGLRAQNKQCSIEWSGSSISKENWIEESKCLGVHIEGFKHWFNFCYCEQ 121
Db 87 SFPHVWIGLDEGKEQCCSSEWSGSSVSVENWIEAESTCLGLELDSNVHKWVNYCGQ 146

QY 122 QDDPFVCEA 129
Db 147 RNPVCEA 154

RESULT 7
Q8UVC6 PRELIMINARY; PRT; 158 AA.
ID Q8UVC6
AC Q8UVC6;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Antithrombin 1 A chain.
OS Agkistrodon acutus (Hundred-pace snake) (Deinagkistrodon acutus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroides;
OC Viperidae; Crotalinae; Deinagkistrodon.
OX NCBI_TaxID=36307;
RN [1]
RP SEQUENCE FROM N.A.
YU Yu H., Xiang K., Liu J.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF463322; AAL66391.1; -.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; Pancreatis_ac.
DR PRINTS; PR01504; PNCREATITSAP.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
SQ SEQUENCE 158 AA; 17950 MW; A73A9C895997BFD7 CRC64;

Query Match 61.4%; Score 439; DB 13; Length 158;
Best Local Similarity 60.2%; Pred. No. 1.4e-38;
Matches 77; Conservative 18; Mismatches 33; Indels 0; Gaps 0;

QY 1 DCSSDSSSYEGHCYKVKQSKTWTDAESFCTKVQNGHLSIESSEADPVGOLIAQKIK 60
Db 26 NCPPGWSAYDQICYQVKEPKNWDDEAEERFCTGAGDGHLSIESKGERDFVAGLVQSIE 85

QY 61 SAKIHVWIGLRAQNKQCSIEWSGSSISKENWIEESKCLGVHIEGFKHWFNFCYCE 120
Db 86 SVEDHVTWTLGVQNKQCSIEWSGSSVSVENWIEAESTCLGLELDSNVHKWVNYCGQ 145

QY 121 QDDPFVCE 128
Db 146 QLNPFVCK 153

```

SEQUENCE F

```

QY 121 QODPFVCE 128
Db 146 RENPFVCK 153

RESULT 11
Q9AV97
ID Q8AV97 PRELIMINARY; PRT; 158 AA.
AC Q8AV97
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Flavocetin-A alpha chain.
OS Trimeresurus flavoviridis (Habu).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;
OC Viperidae; Crotalinae; Trimeresurus.
OX NCBI_TaxID=88087;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RX MEDLINE=20402624; PubMed=10942790;
RA Shin Y., Okuyama I., Hasegawa J., Morita T.;
RT "Molecular cloning of glycoprotein IB-binding protein, flavocetin-A,
RT which inhibits platelet aggregation.";
RL Thromb. Res. 99:239-247(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=venom gland;
RA Shin Y., Okuyama I., Hasegawa J., Morita T.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY149341; AANT72438.1;
SQ SEQUENCE 158 AA, 18073 MW, 8C138650665CA454 CRC64;

Query Match 54.7%; Score 391; DB 13; Length 158;
Best Local Similarity 53.9%; Pred. No. 1.7e-33;
Matches 69; Conservative 21; Mismatches 38; Indels 0; Gaps 0;

QY 1 DCSSDWSYEGHCYKVFQSKTWTDAESFCTQVNGGHLVSISSGEADFGVQLIAQIK 60
Db 26 DCIPGSYDAYDYQAFQKPNKWNDAESFCBEGVKTSHLVSISSGEGDFVAQLVAEKIK 85

QY 61 SAKIHVWIGLRAQNKQCSIEWSDGSSISKENWIEESKCLGVHIETGFHKWENFYCE 120
Db 86 TSFYQVWIGLRIQNKQCSIEWSDGSSISKENWIEESKCLGVHIETGFHKWENFYCE 120

QY 121 QODPFVCE 128
Db 146 TENDPVCK 153

RESULT 12
Q9PSM9
ID Q9PSM9 PRELIMINARY; PRT; 131 AA.
AC Q9PSM9
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE ECLV IX/X-BP alpha SUBUNIT=COAGULATION factor IX/factor X-binding
DE protein alpha subunit.
OS Echis carinatus (Saw-scaled viper).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;
OC Viperidae; Viperinae; Echis.
OX NCBI_TaxID=40353;
RN [1]
RP SEQUENCE.
RX MEDLINE=96196635; PubMed=8611513;
RA Chen Y.L., Teai I.H.;
RT "Functional and sequence characterization of coagulation factor
RT IX/factor X-binding protein from the venom of Echis carinatus
RT leucogaster.";
```

```

RL Biochemistry 35:5264-5271(1996).
DR HSP; P23806; IIXX.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS00441; C_TYPE_LECTIN_2; 1.
SQ SEQUENCE 131 AA, 15439 MW, E85E6C5CBF317E24 CRC64;

Query Match 54.5%; Score 390; DB 13; Length 131;
Best Local Similarity 54.6%; Pred. No. 1.7e-33;
Matches 71; Conservative 17; Mismatches 40; Indels 2; Gaps 2;

QY 1 DCSSDWSYEGHCYKVFQSKTWTDAESFCTQVNGGHLVSISSGEADFGVQLIAQIK- 59
Db 1 DCLPGWSSHEGHCYKVFNEYKTWKDAEFCCKQKSGHLVSVSEEGDFVAKLISENLE 60

QY 60 KSAKI-HWVIGLRAQNKQCSIEWSDGSSISKENWIEESKCLGVHIETGFHKWENFY 118
Db 61 KSHSIDFVWTLTYKGRWKQCSSEWSGSKIKYQKQKQPRKCLGLEKQTEFRKWVNL 120

QY 119 CEQODPFVCE 128
Db 121 CEQODPFVCE 130

RESULT 13
Q9PSM6
ID Q9PSM6 PRELIMINARY; PRT; 142 AA.
AC Q9PSM6
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE Platelet glycoprotein IB-binding protein alpha subunit, GPIB-BP alpha
DE subunit.
OS Bothrops jararaca (Jararaca).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;
OC Viperidae; Crotalinae; Bothrops.
OX NCBI_TaxID=8724;
RN [1]
RP SEQUENCE.
RX MEDLINE=96209990; PubMed=8631868;
RA Kawasaki T., Fujimura Y., Usami Y., Suzuki M., Miura S., Sakurai Y.,
RA Makita K., Taniuchi Y., Hirano K., Titani K.;
RT "Complete amino acid sequence and identification of the platelet
RT glycoprotein IB-binding site of jararaca GPIB-BP, a snake venom
RT protein isolated from Bothrops jararaca.";
RL J. Biol. Chem. 271:10635-10639(1996).
DR HSP; P23806; IIXX.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS00441; C_TYPE_LECTIN_2; 1.
SQ SEQUENCE 142 AA, 16720 MW, E28FDEBEF922004 CRC64;

Query Match 53.5%; Score 382.5; DB 13; Length 142;
Best Local Similarity 51.1%; Pred. No. 1.2e-32;
Matches 68; Conservative 23; Mismatches 37; Indels 5; Gaps 1;

QY 1 DCSSDWSYEGHCYKVFQSKTWT-----DAESFCTQVNGGHLVSISSGEADFGVQLI 55
Db 5 ECPSDWSTHRQYCYKVFQKESNDRSEYDAERFCSEQAKGHLVSISSGEADFGVQLV 64

QY 56 AQIKSAKHVWIGLRAQNKQCSIEWSDGSSISKENWIEESKCLGVHIETGFHKWE 115
Db 65 APNIGSKYVWIGLRIENKQCSSEWSGSKIKYQKQKQPRKCLGLEKQTEFRKWVNL 124

QY 116 NFYCEQODPFVCE 128
Db 125 NIDCVENPFVCK 137
```

DR PROSITE: PS00615; C_TYPE LECTIN 1; 1.
DR PROSITE: PS50041; C_TYPE LECTIN_2; 1.
KW Signal.
FT CHAIN 1 23 POTENTIAL.
SQ SEQUENCE 155 AA; 17798 MW; 92840281A797AACD CRC64;
Query Match 52.8%; Score 377.5; DB 13; Length 155;
Best Local Similarity 54.7%; Pred. No. 4.3e-32;
Matches 70; Conservative 18; Mismatches 39; Indels 1; Gaps 1;
QY 1 DCSSDWSYEGHCYKVFQSKTWTDAESFCTKQVNGHLSIESGSEADFGVGLIAQKIK 60
Db 24 DCLPGWSSYIRFCYQPFKLLKTWEDAERFCTEQANGHLVSPESAREADFGVGLSENK 83
QY 61 SAKIHVWIGLRAQNEKQSIWSDGSSISKENWIEESKCLGVHIETGPHKWFNFCY 120
Db 84 -IKPYVWIGLRYQNEGGQSSKWSKSVENLVPEFSSKCFVLKDTGFTWENVYCG 142
QY 121 QODPFVCE 128
Db 143 LKHVFWCK 150
Search completed: December 8, 2003, 09:54:48
Job time : 64.6571 secs

RESULT 14
Q8JIV6 PRELIMINARY; PRT; 155 AA.
AC Q8JIV6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Antithrombin A A-chain.
OS Agkistrodon acutus (Hundred-pace snake) (Deinagkistrodon acutus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Deinagkistrodon.
OX NCBI_TaxID=36307;
RN [1]
RP SEQUENCE FROM N.A.
RA Yu H., Xiang Y., Wang Y., Liu J.;
RT "A chain of antithrombin A from Deinagkistrodon acutus.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY091762; AAM22790.1; -.
DR InterPro; IPR001304; Lectin C.
DR InterPro; IPR003990; Pancreatins_ac.
DR Pfam; PF00059; lectin c; 1.
DR PRINTS; PR01504; PNCREATITSAP.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE LECTIN 1; 1.
DR PROSITE; PS50041; C_TYPE LECTIN_2; 1.
SQ SEQUENCE 155 AA; 17715 MW; 66E55C95E997A1D7 CRC64;

Query Match 53.2%; Score 380.5; DB 13; Length 155;
Best Local Similarity 54.7%; Pred. No. 2.1e-32;
Matches 70; Conservative 18; Mismatches 39; Indels 1; Gaps 1;
QY 1 DCSSDWSYEGHCYKVFQSKTWTDAESFCTKQVNGHLSIESGSEADFGVGLIAQKIK 60
Db 24 DCLPGWSSYIRFCYQPFKLLKTWEDAERFCTEQANGHLVSPESAREADFGVGLSENK 83
QY 61 SAKIHVWIGLRAQNEKQSIWSDGSSISKENWIEESKCLGVHIETGPHKWFNFCY 120
Db 84 -IKPYVWIGLRYQNEGGQSSKWSKSVENLVPEFSSKCFVLKDTGFTWENVYCG 142
QY 121 QODPFVCE 128
Db 143 LKHVFWCK 150

RESULT 15
Q9DEA2 PRELIMINARY; PRT; 155 AA.
AC Q9DEA2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Agkicetin alpha subunit precursor.
GN GPIBA1.
OS Agkistrodon acutus (Hundred-pace snake) (Deinagkistrodon acutus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Deinagkistrodon.
OX NCBI_TaxID=36307;
RN [1]
RP SEQUENCE FROM N.A.
RA Chen Y.L., Hong T.M., Chang T., Tsai I.H.;
RT "cDNA sequence and functional characterization of glycoprotein Ib-binding protein from the venom of Deinagkistrodon acutus.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF102901; AAG42040.1; -.
DR HSP; P23806; 11XX.
DR InterPro; IPR001304; Lectin C.
DR InterPro; IPR003990; Pancreatins_ac.
DR Pfam; PF00059; lectin c; 1.
DR PRINTS; PR01504; PNCREATITSAP.
DR SMART; SM00034; CLECT; 1.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 8, 2003, 13:34:19 ; Search time 3342.94 Seconds
(without alignments)
937.879 Million cell updates/sec

Title: US-09-938-114-2
Perfect score: 715
Sequence: 1 DCSSWSYEGHCYKVFQKQ.....GPHKWFYCEQDDPFVCEA 129

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US0938114/runat_08122003_091000_22859/app_query.fasta_1.725
-DB=EST_QMWT=fastcap -SURFIX=rest -MINMATCH=0.1 -LOOPECL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptio -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0938114.cgn 1.1 4382 @runat_08122003_091000_22859 -NCPU=6 -ICPU=3
-NO MMAP -LARGOQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: em_estba.*
- 2: em_esthum.*
- 3: em_estin.*
- 4: em_estmu.*
- 5: em_estov.*
- 6: em_estpl.*
- 7: em_estro.*
- 8: em_hic.*
- 9: gb_est1.*
- 10: gb_est2.*
- 11: gb_hic.*
- 12: gb_est3.*
- 13: gb_est4.*
- 14: gb_est5.*
- 15: em_estfun.*
- 16: em_estom.*
- 17: em_gss_hum.*
- 18: em_gss_inv.*
- 19: em_gss_pln.*
- 20: em_gss_vrt.*
- 21: em_gss_fun.*
- 22: em_gss_man.*
- 23: em_gss_mus.*
- 24: em_gss_pro.*
- 25: em_gss_rod.*
- 26: em_gss_phg.*
- 27: em_gss_vri.*
- 28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	382.5	53.5	556	12	BM401668	BM401668 JL2H08F S
2	373.5	52.2	501	12	BM401648	BM401648 JL2B03F S
3	359	50.2	460	12	BM401460	BM401460 JH1D06F S
4	354.5	49.6	618	12	BM401682	BM401682 PH012F Sn
5	342.5	47.9	458	12	BM401641	BM401641 JH1H06F S
6	341.5	47.8	511	12	BM401631	BM401631 JH1F11F S
7	331.5	46.4	406	12	BM401598	BM401598 JH4F11F S
8	324.5	45.4	410	12	BM401605	BM401605 JH4H07F S
9	323	45.2	418	12	BM401662	BM401662 JH2G06F S
10	312.5	43.7	440	12	BM401459	BM401459 JH1D05F S
11	312.5	43.7	453	12	BM401653	BM401653 JH2D01F S
12	312.5	43.7	461	12	BM401683	BM401683 PH012R Sn
13	300.5	42.0	440	12	BM401637	BM401637 JH1G11F S
14	300.5	42.0	456	12	BM401659	BM401659 JH2F01F S
15	279	39.0	654	12	BM401401	BM401401 GH025F Sn
16	277	38.7	574	12	BM401417	BM401417 GH061F Sn
17	248.5	34.8	405	12	BM401607	BM401607 JH4H10F S
18	235	32.9	510	12	BM401667	BM401667 JH2H07F S
19	235	32.9	512	12	BM401495	BM401495 JH2B09F S
20	232.5	32.5	497	12	BM401552	BM401552 JH3E12F S
21	230	32.2	633	12	BM401686	BM401686 PH015FB S
22	226	31.6	504	12	BM401628	BM401628 JH1E12F S
23	223.5	31.3	388	12	BM401625	BM401625 JH1E05F S
24	223.5	31.3	475	12	BM401621	BM401621 JH1D02F S
25	223	31.2	378	12	BM401618	BM401618 JH1C07F S
26	216	30.2	378	12	BM401620	BM401620 JH1C11F S
27	216	30.2	391	12	BM401469	BM401469 JH1E09F S
28	211.5	29.6	365	12	BM401665	BM401665 JH2H01F S
29	210.5	29.4	453	14	CA843095	CA843095 if33d01.Y
30	210.5	29.4	535	14	CA772255	CA772255 i039g12.Y
31	209	29.2	301	12	BM401654	BM401654 JH2G07F S
32	209	29.2	374	12	BM401520	BM401520 JH2G02F S
33	208.5	29.2	486	14	CA773380	CA773380 im63c12.Y
34	208.5	29.2	546	13	BQ776439	BQ776439 i132f09.Y
35	208.5	29.2	642	12	BM312458	BM312458 i978g11.X
36	207.5	29.0	570	14	CA842080	CA842080 i028a05.Y
37	207.5	29.0	628	13	BU072359	BU072359 im45h03.X
38	206.5	28.9	425	12	BM352964	BM352964 i069f02.Y
39	206.5	28.9	551	14	CA775597	CA775597 i087g10.Y
40	206.5	28.9	593	14	CA774835	CA774835 i013b07.X
41	205.5	28.7	453	14	CA868012	CA868012 i176d01.Y
42	205.5	28.7	546	13	BQ778069	BQ778069 i140f01.Y
43	205.5	28.7	547	14	CA772731	CA772731 i084b10.Y
44	205.5	28.7	550	12	BM313297	BM313297 i082d01.Y
45	205.5	28.7	556	12	BM313209	BM313209 i080h07.Y

ALIGNMENTS

RESULT 1
BM401668
LOCUS
DEFINITION
JL2H08F Snake Bothrops insularis library IL3 Bothrops insularis
cDNA 5', similar to Snake venom C-type lectin, mRNA sequence.
ACCESSION
BM401668
VERSION
BM401668.1 GI:20376296
KEYWORDS
EST.
ORGANISM
Bothrops insularis (island jararaca)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
Viperidae; Crotalinae; Bothrops.

```

REFERENCE
AUTHORS
TITLE
1 (bases 1 to 556)
Junqueira-de-Azevedo, I.L.M. and Ho, P.L.
A survey of gene expression and diversity in the venom glands of
the pitviper snake Bothrops insularis through the generation of
expressed sequence tags (ESTs)
Gene 299 (1-2), 279-291 (2002)
22347338
MEDLINE
PUBMED
12459276
COMMENT
Contact: Paulo Lee Ho
Instituto Butantan
Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
Tel: 55 11 37 26 7222 ext. 2083
Fax: 55 11 37 26 1505
Email: hoplee@usp.br
This EST corresponds to cluster BITL07A (see Reference)
Seq primer: M13F.

JOURNAL
MEDLINE
PUBMED
12459276
COMMENT
Contact: Paulo Lee Ho
Instituto Butantan
Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
Tel: 55 11 37 26 7222 ext. 2083
Fax: 55 11 37 26 1505
Email: hoplee@usp.br
This EST corresponds to cluster BITL07A (see Reference)
Seq primer: M13F.

FEATURES
source
Location/Qualifiers
1..556
/organism="Bothrops insularis"
/mol_type="mRNA"
/db_xref="taxon:8723"
/tissue_type="venom glands"
/clone_lib="Snake Bothrops insularis library IL3"
/notes="Organ: venom glands; Vector: pGEM1zf+; Site 1: Eco
RI; Site 2: Not I; Sug of mRNA from Bothrops insularis
venom glands were primed with oligo-(dr) and reverse
transcribed to cDNA using Superscript Plasmid System for
cDNA Synthesis and Cloning (Life Technologies). The cDNAs
were selected by size (350-600 pb and up 600 pb) in
agarose gel electrophoresis, linked to Eco RI adapters and
directionally cloned in pGEM1zf+ vector (Promega). ESTs
were generated from random clones and grouped in unique
sequences. The putative identification of each EST or
cluster was obtained through Blast searches (e-value <
e-05)."

BASE COUNT      146 a   126 c   154 g   129 t       1 others
ORIGIN
Alignment Scores:
Pred. No.:      4.82e-34      Length:      556
Score:          382.50      Matches:      68
Percent Similarity: 68.42%      Conservative: 23
Best Local Similarity: 51.13%      Mismatches: 37
Query Match:     53.50%      Indels:      5
DB:              12      Gaps:        1

US-09-938-114-2 (1-129) x BM401668 (1-556)

QY      1  AspCysSerSerAspTrpSerTyrGluGlyHisCysTyrIysValPheLysGlnSer 20
Db      82  GAGTGTCCCTCTGCTGATGGTCCACCCATAGACAGATGTTCTGTCTCGGAGCGGAAG 141

QY      21  LysThrTrpThr-----AspAlaGluSerPheCysThrLysGlnValAsn 35
Db      142  GAGAGCTGGGATCATGTGACGTGGTATGATGCAGAGAGGTCTGTCTCGGAGCGGAAG 201

QY      36  GlyGlyHisLeuValSerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIle 55
Db      202  GCGCGCCATCTGCTCTATCCAAAGCGATGAAGAAGACAGACTTTGTGGCCAGCTGGTC 261

QY      56  AlaGlnIleLysSerAlaIleHisValTrpIleGlyLeuArgAlaGlnAsnLys 75
Db      262  GCTCCGAACATCGGGAATCAATACTACTGTCTGGATCGGATCGAGATTGAAACAAA 321

QY      76  GluLysGlnCysSerIleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIle 95
Db      322  AAACAGCAATCGAGCTCGAGTGGAGCGATTACTCCAGCGTCAAGTTATGAGAACCTGGTT 381

QY      96  GluGluGluSerLysLysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGlu 115
Db      382  CGAGGAATGTGAAAAAGTGTTCGCTCGCAAAAAAAGCAAGGGTTTCGTAAGTGGTTC 441

```

```

QY      116  AsnPheTyrCysGluGlnGlnAspPropheValCysGlu 128
Db      442  AATATTGACTGTGTAGAGGAATCCTTCTGTGTGCAAG 480

RESULT 2
LOCUS      BM401648
DEFINITION  BM401648 Snake Bothrops insularis library IL3 Bothrops insularis
            cDNA 5' similar to Snake venom C-type lectin, mRNA sequence.
ACCESSION  BM401648
VERSION    BM401648.1 GI:20376276
KEYWORDS   EST.
SOURCE     Bothrops insularis (island jararaca)
ORGANISM   Bothrops insularis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;
            Viperidae; Crotalinae; Bothrops.
            1 (bases 1 to 501)
            Junqueira-de-Azevedo, I.L.M. and Ho, P.L.
            A survey of gene expression and diversity in the venom glands of
            the pitviper snake Bothrops insularis through the generation of
            expressed sequence tags (ESTs)
Gene 299 (1-2), 279-291 (2002)
22347338
MEDLINE
PUBMED
12459276
COMMENT
Contact: Paulo Lee Ho
Centro de Biotecnologia
Instituto Butantan
Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
Tel: 55 11 37 26 7222 ext. 2083
Fax: 55 11 37 26 1505
Email: hoplee@usp.br
This EST corresponds to cluster BITL07A (see Reference)
Seq primer: M13F.
Location/Qualifiers
1..501
/organism="Bothrops insularis"
/mol_type="mRNA"
/db_xref="taxon:8723"
/tissue_type="venom glands"
/clone_lib="Snake Bothrops insularis library IL3"
/notes="Organ: venom glands; Vector: pGEM1zf+; Site 1: Eco
RI; Site 2: Not I; Sug of mRNA from Bothrops insularis
venom glands were primed with oligo-(dr) and reverse
transcribed to cDNA using Superscript Plasmid System for
cDNA Synthesis and Cloning (Life Technologies). The cDNAs
were selected by size (350-600 pb and up 600 pb) in
agarose gel electrophoresis, linked to Eco RI adapters and
directionally cloned in pGEM1zf+ vector (Promega). ESTs
were generated from random clones and grouped in unique
sequences. The putative identification of each EST or
cluster was obtained through Blast searches (e-value <
e-05)."

BASE COUNT      134 a   102 c   139 g   122 t       4 others
ORIGIN
Alignment Scores:
Pred. No.:      4.55e-33      Length:      501
Score:          373.50      Matches:      67
Percent Similarity: 66.42%      Conservative: 22
Best Local Similarity: 50.00%      Mismatches: 40
Query Match:     52.24%      Indels:      5
DB:              12      Gaps:        1

US-09-938-114-2 (1-129) x BM401648 (1-501)

QY      1  AspCysSerSerAspTrpSerTyrGluGlyHisCysTyrIysValPheLysGlnSer 20
Db      84  GAGTGTCCCTCTGCTGATGGTCCACCCATAGACAGATGTTCTGTCTCGGAGCGGAAG 143

QY      21  LysThrTrpThr-----AspAlaGluSerPheCysThrLysGlnValAsn 35
Db      144  GAGAGCTGGGATCATGTGACGTGGTATGATGCAGAGAGGTCTGTCTCGGAGCGGAAG 203

```

QY 36 GlyClyHisLeuValSerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIle 55
 Db 204 GCGCGCCATCTGGTCTCTATCGAAGCGATGAAGACGACACTTGTGGCCCACTGGTC 263
 QY 56 AlaGlnLysIleLysSerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLys 75
 Db 264 GCTCCGAACATAGGGAATCAATATCTATGTCGGATCGGACTGAGGATTGAACACAA 323
 QY 76 GluLysGlnCysSerIleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIle 95
 Db 324 AAACAGCAATCGAGCTCGAAGTGGAGGATTTACTCCAGGGTCAGTTATGAGAACCTGGT 383
 QY 96 GluGluGluSerLysLysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGlu 115
 Db 384 CGAGGAATGTGAAGAGTGTTCCTGCGTGAAGAAAGCAAGGGTTTCNTAAGTGGTC 443
 QY 116 AsnPheTyrcysGluGlnGlnAspPheValCysGluAla 129
 Db 444 AATATTGACTGTGTANAGGAATCTTCTTGTGCAATTCA 485

RESULT 3
 BM401460 460 bp mRNA linear EST 01-MAY-2002
 LOCUS JH1D06F Snake Bothrops insularis library IL3 Bothrops insularis
 DEFINITION cDNA 5', similar to Snake venom C-type lectin, mRNA sequence.
 ACCESSION BM401460.1 GI:20376088
 VERSION
 KEYWORDS EST.
 SOURCE Bothrops insularis (island jararaca)
 ORGANISM Bothrops insularis

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 TITLE Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
 Viperidae; Crotalinae; Bothrops.
 JOURNAL Junqueira-de-Azevedo, I.L.M. and Ho, P.L.
 MEDLINE A survey of gene expression and diversity in the venom glands of
 PUBMED the pitviper snake Bothrops insularis through the generation of
 COMMENT expressed sequence tags (ESTs)
 Gene 299 (1-2), 279-291 (2002)
 12459276
 CONTACT Paulo Lee Ho
 INSTITUTION Centro de Biotecnologia
 AV. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
 TEL: 55 11 37 26 7222 ext. 2083
 FAX: 55 11 37 26 1505
 EMAIL: hoplee@usp.br
 THIS EST corresponds to cluster BITL13A (see Reference)
 SEQ PRIMER: M13P.

FEATURES
 source

1. 460
 /organism="Bothrops insularis"
 /mol_type="mRNA"
 /db_xref="taxon:8723"
 /tissue_type="venom glands"

/clone_lib="Snake Bothrops insularis library IL3"
 /note="Organ: venom glands; Vector: pGEM11zf+; Site 1: Eco
 RI; Site 2: Not I; Sug of mRNA from Bothrops insularis
 venom glands were primed with oligo-(dT) and reverse
 transcribed to cDNA using Superscript Plasmid System for
 cDNA Synthesis and Cloning (Life Technologies). The cDNAs
 were selected by size (350-600 pb and up 600 pb) in
 agarose gel electrophoresis, linked to Eco RI adapters and
 directionally cloned in pGEM11zf+ vector (Promega). ESTs
 were generated from random clones and grouped in unique
 sequences. The putative identification of each EST or
 cluster was obtained through Blast searches (e-value <
 e-05)."

106 a 106 c 145 g 103 t

BASE COUNT
 ORIGIN

Alignment Scores:

Pred. No.: 1.91e-31 Length: 460
 Score: 359.00 Matches: 61
 Percent Similarity: 74.51% Conservative: 15
 Best Local Similarity: 59.80% Mismatches: 26
 Query Match: 50.21% Indels: 0
 DB: 12 Gaps: 0

US-09-938-114-2 (1-129) x BM401460 (1-460)

QY 1 AspCysSerSerAspTrpSerSerTrpGluClyHisCysTrpLysValPheLysGlnSer 20
 Db 143 GATTGTCCTCCCTGATTGCTCTCTATGAAGGAGTTGCTACAGGGTCTTCACTGAACCG 202
 QY 21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
 Db 203 CAAAACTGGGCGATGTCAGAGAGTTCTCTCGAGCAGCGAAGCGCGCCATCTGCTG 262
 QY 41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
 Db 263 TCTATCGAAAGCGAAGGAGAGCAGACTTTGTGCCCGCAGCTGCTGAGAGGATAGAC 322
 QY 61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
 Db 323 AGACCCCAATCCCATGTTCTGATCGGACTGAGGGGTGACACAAAGGACGCAATGAGC 382
 QY 81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluSerLys 100
 Db 383 TCGAACTGGAGCGATGCTCTGCTCTGTTATGAGAACCTGTTGAACGAAGAACGAAA 442
 QY 101 LysCys 102
 Db 443 AATGT 448

RESULT 4

BM401682

LOCUS

DEFINITION

PH012F Snake Bothrops insularis library IL2 Bothrops insularis cDNA

5', similar to Snake venom C-type lectin, mRNA sequence.

ACCESSION

BM401682

VERSION

BM401682.1 GI:20376310

KEYWORDS EST.

SOURCE

Bothrops insularis (island jararaca)

Bothrops insularis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;

Viperidae; Crotalinae; Bothrops.

REFERENCE

1 (bases 1 to 618)

A survey of gene expression and diversity in the venom glands of

the pitviper snake Bothrops insularis through the generation of

expressed sequence tags (ESTs)

Gene 299 (1-2), 279-291 (2002)

22347338

PUBMED

COMMENT

Contact: Paulo Lee Ho

Centro de Biotecnologia

Instituto Butantan

Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900

Tel: 55 11 37 26 7222 ext. 2083

Fax: 55 11 37 26 1505

Email: hoplee@usp.br

This EST corresponds to cluster BITL09A (see Reference)

Seq primer: M13P.

Location/Qualifiers

1. 618

/organism="Bothrops insularis"

/mol_type="mRNA"

/db_xref="taxon:8723"

/tissue_type="venom glands"

/clone_lib="Snake Bothrops insularis library IL2"

/note="Organ: venom glands; Vector: pGEM11zf+; Site 1: Eco

RI; Site 2: Not I; Sug of mRNA from Bothrops insularis

venom glands were primed with oligo-(dT) and reverse transcribed to cDNA using Superscript Plasmid System for cDNA Synthesis and Cloning (Life Technologies). The cDNAs were selected by size (350-600 pb and up 600 pb) in agarose gel electrophoresis, linked to Eco RI adapters and directionally cloned in pGEM11zf+ vector (Promega). ESTs were generated from random clones and grouped in unique sequences. The putative identification of each EST or cluster was obtained through Blast searches (e-value < e-05).

BASE COUNT 157 a 124 c 180 g 149 t 8 others
ORIGIN

Alignment Scores:
Pred. No.: 9.71e-31 Length: 618
Score: 354.50 Matches: 72
Percent Similarity: 71.09% Conservatives: 19
Best Local Similarity: 56.25% Mismatches: 36
Query Match: 49.58% Indels: 4
DB: 12 Gaps: 1

US-09-938-114-2 (1-129) x BM401682 (1-618)

Qy 1 AspCysSerSerAspTyrSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
Db 214 GATTGTAAACCTTATTGGTCTCTCTATGAAGGGCGTCCACAGGTCTTCAAGAAGCG 273
Qy 21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAenGlyGlyHisLeuVal 40
Db 274 AAAAACTGGGACGATGACAGAGGTTCTGCTCGGAGGAGTGAATGGCGGCATCTGGTC 333
Qy 41 SerileGluSerSerGlyGluAlaAspPheValGlnLeuIleAlaGlnLysIleLys 60
Db 334 TCTATCAAGACCGCGGAGAGACGCTTTGTGGCCAGCTGGTGGCTGGAACAACAC 393
Qy 61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAenLysGluLysGlnCysSer 80
Db 394 AGTCGCAATGAATGTCGTGATCGAATGAGGTTCAAGGCAAGA-AAACAATGCANC 452
Qy 81 IleGluTrpSerAspGlySerSerIleSerLysGluAenTrpIleGluGluGluSerLys 100
Db 453 ---TCGTGGGCGATGGTCCAC-GTCAGTTATGAGAACTGATTAAGCANAATTA 508
Qy 101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAenPheTyrCysGlu 120
Db 509 ACATGATGTGGTGTGA AAAAAGATGCGAGGTTTNTAAGTGAGCAATCTTAACGTGAA 568
Qy 121 GlnGlnAspPropheValCysGlu 128
Db 569 CAACNAATTCCTTCTC-TGGAG 591

RESULT 5

BM401641

LOCUS

DEFINITION BM401641 458 bp mRNA linear EST 01-MAY-2002
JL1H06F Snake Bothrops insularis library IL3 Bothrops insularis
cDNA 5' similar to Snake venom C-type lectin, mRNA sequence.

ACCESSION BM401641

VERSION BM401641.1 GI:20376269

KEYWORDS EST.

SOURCE Bothrops insularis (island jararaca)

ORGANISM Bothrops insularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Lepidodermata; Squamata; Scleroglossa; Serpentes; Colubroidea;
Viperidae; Crotalinae; Bothrops.
1 (bases 1 to 458)

REFERENCE Junqueira-de-Azevedo, I.L.M. and Ho, P.L.

A survey of gene expression and diversity in the venom glands of
the pitviper snake Bothrops insularis through the generation of
expressed sequence tags (ESTs)

Gene 299 (1-2), 279-291 (2002)

JOURNAL 22347338

MEDLINE 12459276

COMMENT Contact: Paulo Lee Ho

Centro de Biotecnologia
Instituto Butantan
Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
Tel: 55 11 37 26 7222 ext. 2083
Fax: 55 11 37 26 1505
Email: hoplee@usp.br
This EST corresponds to cluster BITL07A (see Reference)
Seq primer: M13F.

FEATURES

Location/Qualifiers

1..458

/organism="Bothrops insularis"

/mol_type="mRNA"

/db_xref="taxon:8723"

/tissue_type="venom glands"

/clone_lib="Snake Bothrops insularis library IL3"

/note="Organ: venom glands; Vector: pGEM11zf+; Site 1: Eco

RI; Site 2: Not I; Sug of mRNA from Bothrops insularis

venom glands were primed with oligo-(dT) and reverse

transcribed to cDNA using Superscript Plasmid System for

cDNA Synthesis and Cloning (Life Technologies). The cDNAs

were selected by size (350-600 pb and up 600 pb) in

agarose gel electrophoresis, linked to Eco RI adapters and

directionally cloned in pGEM11zf+ vector (Promega). ESTs

were generated from random clones and grouped in unique

sequences. The putative identification of each EST or

cluster was obtained through Blast searches (e-value <

e-05).

BASE COUNT 122 a 90 c 130 g 111 t 5 others

ORIGIN

Alignment Scores:

Pred. No.: 1.53e-29 Length: 458
Score: 342.50 Matches: 62
Percent Similarity: 66.13% Conservatives: 20
Best Local Similarity: 50.00% Mismatches: 37
Query Match: 47.90% Indels: 5
DB: 12 Gaps: 1

US-09-938-114-2 (1-129) x BM401641 (1-458)

Qy 1 AspCysSerSerAspTyrSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20

Db 84 GAGTGTCCCTCTGATTGGTCCACCACAGACAGATTGTACAAAGTTCTTCCACAAAG 143

Qy 21 LysThrTrpThr-----AspAlaGluSerPheCysThrLysGlnValasn 35

Db 144 GAGAGCTGGGATGATGTGAGCTGTATGATGCANANAGTTCTGCTCGGACGCGAAG 203

Qy 36 GlyGlyHisLeuValSerIleGluSerSerGlyGluAlaAspPheValGlyLeuIle 55

Db 204 GCGGCCCATCTGGTCTCTATCGAAGCGATGAAGACAGACTTTTGGCCACGCTGGTC 263

Qy 56 AlaGlnLysIleLysSerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLys 75

Db 264 GCTCCGAACATAGGAAATCCAAATACTATGCTCGATCGGACTCGAGATTGAAACAAA 323

Qy 76 GluLysGlnCysSerIleGluTrpSerAspGlySerIleSerLysGluAsnTrpIle 95

Db 324 AACAGCAATGACGTGAGTGGAGGATTTACTTCAGCGGTTCAGTTATGAGAACTGGTT 383

Qy 96 GluGluGluSerLysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGlu 115

Db 384 CGAGGAATGTGAAAAGTGTTCGTCGGAAGAAAGCAANGTTTCNTAAGTGGTC 443

Qy 116 AsnPheTyrCys 119

Db 444 AATATTGACTGT 455

RESULT 6

BM401631

LOCUS

DEFINITION BM401631 511 bp mRNA linear EST 01-MAY-2002

JL1F11F Snake Bothrops insularis library IL3 Bothrops insularis

cDNA 5' similar to Snake venom C-type lectin, mRNA sequence.

QY	76	GlUgLnCysSerIleGluTrpSerAspGlySerIleSerIleSerLysGluAsnTrpIle 95
Db	388	AAACANCAATGAGCTCGAAGTGGAGGATCTCCAGCGTCAGTTATGAGAACCTGGT 447
QY	96	GlUgLnGluSerLysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGlu 115
Db	448	CGAGGAAATGTGAAAAAGTGTTCGCTCGGAAAAAGCAAGGGTTCNTAAGTGGGTC 507
QY	116	Asn 116
Db	508	AAAT 510
RESULT 7		
LOCUS	BM401598	406 bp mRNA linear EST 01-MAY-2002
DEFINITION	JH4Fl1F Snake Bothrops insularis library IL3 Bothrops insularis	
ACCESSION	BM401598	
VERSION	BM401598.1	GI:20376226
KEYWORDS	EST.	
SOURCE	Bothrops insularis (island jararaca)	
ORGANISM	Bothrops insularis	
REFERENCE	1 (bases 1 to 406)	
AUTHORS	Junqueira-de-Azevedo, I.L.M. and Ho, P.L.	
TITLE	A survey of gene expression and diversity in the venom glands of the pitviper snake Bothrops insularis through the generation of expressed sequence tags (ESTs)	
JOURNAL	Gene 299 (1-2), 279-291 (2002)	
MEDLINE	22347338	
PUBMED	12459276	
COMMENT	Contact: Paulo Lee Ho Centro de Biotecnologia Instituto Butantan Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900 Tel: 55 11 37 26 7222 ext. 2083 Fax: 55 11 37 26 1505 Email: hoplee@usp.br This EST corresponds to cluster BITL07A (see Reference) Seq primer: M13F.	
FEATURES	Location/Qualifiers	
source	1..406	
	/organism="Bothrops insularis"	
	/mol_type="mRNA"	
	/db_xref="taxon:8723"	
	/tissue_type="venom glands"	
	/clone_lib="Snake Bothrops insularis library IL3"	
	/note="Organ: venom glands; Vector: pGEM11zf+; Site 1: Eco RI; Site 2: Not I; Sug of mRNA from Bothrops insularis venom glands were primed with oligo-(dT) and reverse transcribed to cDNA using Superscript Plasmid System for cDNA Synthesis and Cloning (Life Technologies). The cDNAs were selected by size (350-600 pb and up 600 pb) in agarose gel electrophoresis, linked to Eco RI adapters and directionally cloned in pGEM11zf+ vector (Promega). ESTs were generated from random clones and grouped in unique sequences. The putative identification of each EST or cluster was obtained through Blast searches (e-value < e-05)."	
BASE COUNT	95 a 91 c 126 g 94 t	
ORIGIN		
Alignment Scores:		
Pred. No.:	2,41e-28	Length: 406
Score:	331.50	Matches: 59
Percent Similarity:	78.49%	Conservative: 14
Best Local Similarity:	63.44%	Mismatches: 19
Query Match:	46.36%	Indels: 1
DB:	12	Gaps: 1
US-09-938-114-2 (1-129) x BM401598 (1-406)		

```

Qy 1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
Db 129 GATTGTAACCCCTATTGCTCTCTATGAGGGGGTTCCTACAGGTTCCTTCAAAACACCG 188
Qy 21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
Db 189 AAAAAGCTGGGACGATGACAGAGGTTCTGCTGGAGGAGGTGAATGGCGGCATCTGTC 248
Qy 41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
Db 249 TCTATCGAAGCGCGGAGAGACAGACTTCTGCGCCAGCTGGTCACTGAGAACATACAG 308
Qy 61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
Db 309 AGCCGCGAATGTTGTTGTCGATCGAATGAAGGTTCAAGGCAAGAAAGCAATGTCAGC 368
Qy 81 IleGluTrpSerAspGlySerSerIleSerLysGluAsn 93
Db 369 ---TCGTGGAGCGATGGCTCCAGCGTCAGTTATGAGAAC 404

RESULT 8
BM401605
LOCUS
DEFINITION
JHAH07F Snake Bothrops insularis library IL3 Bothrops insularis
CDNA 5' similar to Snake venom C-type lectin, mRNA sequence.
ACCESSION
BM401605.1 GI:20376233
VERSION
EST.
KEYWORDS
SOURCE
Bothrops insularis (island jararaca)
ORGANISM
Bothrops insularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
Viperidae; Crotalinae; Bothrops.
REFERENCE
1 (bases 1 to 410)
Junqueira-de-Azevedo, I.L.M. and Ho, P.L.
A survey of gene expression and diversity in the venom glands of
the pitviper snake Bothrops insularis through the generation of
expressed sequence tags (ESTs)
Gene 299 (1-2), 279-291 (2002)
22347338
12459276
PUBMED
COMMENT
Contact: Paulo Lee Ho
Centro de Biotecnologia
Instituto Butantan
Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
Tel: 55 11 37 26 7222 ext. 2083
Fax: 55 11 37 26 1505
Email: hoplees@usp.br
This EST corresponds to cluster BITL07A (see Reference)
Seq primer: M13F.
Location/Qualifiers
1..410
/organism="Bothrops insularis"
/mol_type="mRNA"
/db_xref="taxon:8723"
/tissue_type="venom glands"
/clone_lib="Snake Bothrops insularis library IL3"
/note="Organ: venom glands; Vector: pGEMilzf+; Site 1: Eco
RI; Site 2: Not I; 5'ug of mRNA from Bothrops insularis
venom glands were primed with oligo-(dT) and reverse
transcribed to cDNA using Superscript Plasmid System for
cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs
were selected by size (350-600 pb and up 600 pb) in
agarose gel electrophoresis, linked to Eco RI adapters and
directionally cloned in pGEMilzf+ vector (Promega). ESTs
were generated from random clones and grouped in unique
sequences. The putative identification of each EST or
cluster was obtained through Blast searches (e-value <
e-05)."
106 a 87 c 120 g 97 t
BASE COUNT
ORIGIN

```

Alignment Scores:

```

Pred. No.: 1.58e-27 Length: 410
Score: 324.50 Matches: 59
Percent Similarity: 71.30% Conservative: 18
Best Local Similarity: 54.63% Mismatches: 26
Query Match: 45.38% Indels: 5
DB: 12 Gaps: 1

```

US-09-938-114-2 (1-129) x BM401605 (1-410)

```

Qy 1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
Db 84 GAGTGTCCCTCTGATTGGTCCACCACATAGACAGATTGCTACAAAGTCTTCCACAAAG 143
Qy 21 LysThrTrpThr-----AspAlaGluSerPheCysThrLysGlnValAsn 35
Db 144 GAGAGCTGGGATGATGTCAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 203
Qy 36 GlyGlyHisLeuValSerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIle 55
Db 204 GCGGCCATCTGGTCTCTATCGAAGCGATGAAGAGCAGACTTTGTGGCCAGCTGTC 263
Qy 56 AlaGlnLysIleLysSerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLys 75
Db 264 GCTCGAACATAGCGAAATCCAAATACTATGCTGATCGGACTCGGACTGAGGATTGAAACAA 323
Qy 76 GluLysGlnCysSerIleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIle 95
Db 324 AACAGCAATCGAGCTCGAAGTGGAGCGATTACTCCAGCGTCAGTTATGAGAACCTGGTT 383
Qy 96 GluGluGluSerLysLysCysLeu 103
Db 384 CGAGCAATGCGAAAGTGTGTTG 407

```

RESULT 9

BM401662

LOCUS

DEFINITION

CDNA 5' similar to Snake venom C-type lectin, mRNA sequence.

ACCESSION

BM401662

VERSION

BM401662.1 GI:20376290

KEYWORDS

EST.

SOURCE

Bothrops insularis (island jararaca)

ORGANISM

Bothrops insularis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;

Viperidae; Crotalinae; Bothrops.

REFERENCE

1 (bases 1 to 418)

Junqueira-de-Azevedo, I.L.M. and Ho, P.L.

A survey of gene expression and diversity in the venom glands of

the pitviper snake Bothrops insularis through the generation of

expressed sequence tags (ESTs)

Gene 299 (1-2), 279-291 (2002)

22347338

12459276

PUBMED

COMMENT

Contact: Paulo Lee Ho

Centro de Biotecnologia

Instituto Butantan

Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900

Tel: 55 11 37 26 7222 ext. 2083

Fax: 55 11 37 26 1505

Email: hoplees@usp.br

This EST corresponds to cluster BITL07A (see Reference)

Seq primer: M13F.

Location/Qualifiers

1..418

/organism="Bothrops insularis"

/mol_type="mRNA"

/db_xref="taxon:8723"

/tissue_type="venom glands"

/clone_lib="Snake Bothrops insularis library IL3"

/note="Organ: venom glands; Vector: pGEMilzf+; Site 1: Eco

RI; Site 2: Not I; 5'ug of mRNA from Bothrops insularis

venom glands were primed with oligo-(dT) and reverse

transcribed to cDNA using Superscript Plasmid System for

cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs

were selected by size (350-600 pb and up 600 pb) in

agarose gel electrophoresis, linked to Eco RI adapters and

directionally cloned in pGEMilzf+ vector (Promega). ESTs

were generated from random clones and grouped in unique

sequences. The putative identification of each EST or

cluster was obtained through Blast searches (e-value <

e-05)."

FEATURES

source

1..418

/organism="Bothrops insularis"

/mol_type="mRNA"

/db_xref="taxon:8723"

/tissue_type="venom glands"

/clone_lib="Snake Bothrops insularis library IL3"

/note="Organ: venom glands; Vector: pGEMilzf+; Site 1: Eco

RI; Site 2: Not I; 5'ug of mRNA from Bothrops insularis

venom glands were primed with oligo-(dT) and reverse

transcribed to cDNA using Superscript Plasmid System for

cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs

were selected by size (350-600 pb and up 600 pb) in

agarose gel electrophoresis, linked to Eco RI adapters and

directionally cloned in pGEMilzf+ vector (Promega). ESTs

were generated from random clones and grouped in unique

sequences. The putative identification of each EST or

cluster was obtained through Blast searches (e-value <

e-05)."

venom glands were primed with oligo-(dT) and reverse transcribed to cDNA using Superscript Plasmid System for cDNA Synthesis and Cloning (Life Technologies). The cDNAs were selected by size (350-600 pb and up 600 pb) in agarose gel electrophoresis, linked to Eco RI adapters and directionally cloned in pGEM11zf+ vector (Promega). ESTs were generated from random clones and grouped in unique sequences. The putative identification of each EST or cluster was obtained through Blast searches (e-value < e-05)."

BASE COUNT 116 a 90 c 122 g 90 t

Alignment Scores:

Pred. No.: 2,42e-27 Length: 418
Score: 323.00 Matches: 58
Percent Similarity: 70.09% Conservativeness: 17
Best Local Similarity: 54.21% Mismatches: 32
Query Match: 45.17% Indels: 0
DB: 12 Gaps: 0

US-09-938-114-2 (1-129) x BM401662 (1-418)

Qy 22 ThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyLysLeuValSer 41
Db 14 AGCTGGTATGATGCAGAGAGTTCTGCTCGAGCAGCGAGGCGGCACTGCTCTCT 73
Qy 42 IleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysSer 61
Db 74 ATCGAAAGCGATGAAGAAGCAGACTTTGTGGCCAGCTGGTCTCGCAACATAGGAA 133
Qy 62 AlaLysIleHisValTrpIleGlyLeuArgLaglnAsnLysGlyLysGlnCysSerIle 81
Db 134 TCCAAATACTATGTCGATCGACTGAGGATTCGAAACAAAAACACAGCAATGCGCTG 193
Qy 82 GluTrpSerAspGlySerSerIleSerLysGlnValAsnTrpIleGluGluSerLysLys 101
Db 194 AAGTGGAGCGATTAATCTCAGCGTCAGTATGATGAGAACCTGGTTCGAGGAATGTGAAGA 253
Qy 102 CysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGluGln 121
Db 254 TGTTTTCGCTGGAAAAAAGCAAGGGTTTCGTAAGTGGGTCAATATTGCTGTGTAGAA 313
Qy 122 GlnAspProPheValCysGlu 128
Db 314 GGAATCTCTTCGTGTGCAAG 334

RESULT 10

BM401459 440 bp mRNA linear EST 01-MAY-2002
LOCUS JH1D05F Snake Bothrops insularis library IL3 Bothrops insularis
DEFINITION cDNA 5' similar to Snake venom C-type lectin, mRNA sequence.

BM401459

BM401459.1 GI:20376087

EST.

Bothrops insularis (island jararaca)

Bothrops insularis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;

Viperidae; Crotalinae; Bothrops.

Junqueira-de-Azevedo, I.L.M. and Ho, P.L.

A survey of gene expression and diversity in the venom glands of

the pitviper snake Bothrops insularis through the generation of

expressed sequence tags (ESTs)

Gene 299 (1-2), 279-291 (2002)

22347338

12459276

COMMENT

Contact: Paulo Lee Ho

Centro de Biotecnologia

Instituto Butantan

Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900

Tel: 55 11 37 26 7222 ext. 2083

Fax: 55 11 37 26 1505

Email: hoplee@usp.br

This EST corresponds to cluster BITL07A (see Reference)

Seq primer: M13P.

FEATURES

source

Location/Qualifiers

1..440

/organism="Bothrops insularis"

/mol_type="mRNA"

/db_xref="taxon:8723"

/tissue_type="venom glands"

/clone_lib="Snake Bothrops insularis library IL3"

/note="Organ: venom glands; Vector: pGEM11zf+; Site 1: Eco RI; Site 2: Not I; Sug of mRNA from Bothrops insularis venom glands were primed with oligo-(dT) and reverse transcribed to cDNA using Superscript Plasmid System for cDNA Synthesis and Cloning (Life Technologies). The cDNAs were selected by size (350-600 pb and up 600 pb) in agarose gel electrophoresis, linked to Eco RI adapters and directionally cloned in pGEM11zf+ vector (Promega). ESTs were generated from random clones and grouped in unique sequences. The putative identification of each EST or cluster was obtained through Blast searches (e-value < e-05)."

BASE COUNT 113 a 96 c 130 g 101 t

ORIGIN

Alignment Scores:

Pred. No.: 4,26e-26 Length: 440
Score: 312.50 Matches: 58
Percent Similarity: 72.82% Conservativeness: 17
Best Local Similarity: 56.31% Mismatches: 22
Query Match: 43.71% Indels: 6
DB: 12 Gaps: 1

US-09-938-114-2 (1-129) x BM401459 (1-440)

Qy 1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
Db 134 GAGTGTCCCTCTGATTGCTCCACCATAGACAGTATTGCTTCTTCAACAAAG 193
Qy 21 LysThrTrpThr-----AspAlaGluSerPheCysThrLysGlnValAsn 35
Db 194 GAGAGCTGGATGATGATGCTGGTATGTCAGAGAGTTCTCTCGAGCAGCGGAG 253
Qy 36 GlyGlyHisLeuValSerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIle 55
Db 254 GCGGCGCATCTGCTCTCTATCGAAGCGATCAAGACAGACACTTTGGCCAGCTGGTC 313
Qy 56 AlaGlnLysIleLysSerAlaLysIleHisValTrpIleGlyLeuA-gAlaGlnAsnLys 75
Db 314 GCTCCGAACATAGGGAATACTATGCTCGATCGAGCTGAGGATTGAAACAAA 373
Qy 76 GluLysGlnCysSerIleGluTrpSerAspClySerSerIleSerLysGlnAsn-TrpIle 95
Db 374 AACAGCAATCGAGCTCGAAGTGAGCGGATTAATCCAGCGTCAGTTATGAGACCTGTT 433
Qy 95 eGluGlu 97
Db 434 CGAGGAA 440

RESULT 11

BM401653

LOCUS

DEFINITION

BM401653

ACCSSION

BM401653

VERSION

BM401653.1 GI:20376281

KEYWORDS

SOURCE

ORGANISM

Bothrops insularis (island jararaca)

Bothrops insularis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;

Viperidae; Crotalinae; Bothrops.

BM401653 453 bp mRNA linear EST 01-MAY-2002

JL2D01F Snake Bothrops insularis library IL3 Bothrops insularis

cDNA 5' similar to Snake venom C-type lectin, mRNA sequence.

BM401653

BM401653.1 GI:20376281

EST.

Bothrops insularis (island jararaca)

Bothrops insularis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;

Viperidae; Crotalinae; Bothrops.

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

1 (bases 1 to 453)
Junqueira-de-Azevedo, I.L.M. and Ho, P.L.
A survey of gene expression and diversity in the venom glands of
the pitviper snake Bothrops insularis through the generation of
expressed sequence tags (ESTs)
Gene 299 (1-2), 279-291 (2002)
22347338
12459276
Contact: Paulo Lee Ho
Centro de Biotecnologia
Instituto Butantan
Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
Tel: 55 11 37 26 7222 ext. 2083
Fax: 55 11 37 26 1505
Email: hople@usp.br
This EST corresponds to cluster BITL07A (see Reference)
Seq primer: M13F.

FEATURES
source
Location/Qualifiers
1. 453
/organism="Bothrops insularis"
/mol_type="mRNA"
/db_xref="taxon:8723"
/tissue type="venom glands"
/clone lib="Snake Bothrops insularis library IL3"
/note="Organ: venom glands; Vector: pGEM11zf+; Site 1: Eco
RI; Site 2: Not I; 5'ug of mRNA from Bothrops insularis
venom glands were primed with oligo-(dT) and reverse
transcribed to cDNA using Superscript Plasmid System for
cDNA Synthesis and Cloning (Life Technologies). The cDNAs
were selected by size (350-600 pb and up 600 pb) in
agarose gel electrophoresis, linked to Eco RI adapters and
directionally cloned in pGEM11zf+ vector (Promega). ESTs
were generated from random clones and grouped in unique
sequences. The putative identification of each EST or
cluster was obtained through Blast searches (e-value <
e-05)."

BASE COUNT 120 a 96 c 132 g 104 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 453
Score: 312.50
Percent Similarity: 72.81%
Best Local Similarity: 56.31%
Query Match: 43.71%
DB: 12
Length: 453
Matches: 58
Conservative: 17
Mismatch: 22
Indels: 6
Gaps: 1

US-09-938-114-2 (1-129) x BM401653 (1-453)

QY 1 AspCysSerSerAspTyrSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
DB 135 GAGTGTCCCTCTGATGGTCCACCATAGACAGTATGCTACAGTTCTTCCAAACAAAG 194
QY 21 LysThrTrpThr-----AspAlaGluSerPheCysThrLysGlnValAsn 35
DB 195 GAGAGCTGGGATGATGAGTGGTGTATGATGACAGAGGTTCTGCTCGGAGCAGCGAAG 254
QY 36 GlyGlyHisLeuValSerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeu 55
DB 255 GCGCGCCATCTGTCTCTATCCAAAGCGATGAGAGCAGACTTTGTGGCCAGCTGGTTC 314
QY 56 AlaGlnLysIleLysSerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLys 75
DB 315 GCTCCNAACATAGGGAATCCAAATACATATGCTGGATCGGACTGAGGATGAAACAAA 374
QY 76 GluLysGlnCysSerIleGluTrpSerAspGlySerSerIleSerLysGlnAsn-TrpI 95
DB 375 AAACAGCATGCGAGTCCGAAGTGGAGCGATTACTCCAGCGTCAGTTATGAGAACCTGGTT 434
QY 95 eGluGlu 97
DB 435 CGAGGAA 441

RESULT 12
BM401683/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

BM401683
PH012R Snake Bothrops insularis library IL2 Bothrops insularis cDNA
3', similar to Snake venom C-type lectin, mRNA sequence.
BM401683
BM401683.1 GI:20376311
EST.
Bothrops insularis (island jararaca)
Bothrops insularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Lepidodactylia; Squamata; Scleroglossa; Serpentes; Colubroidae;
Viperidae; Crotalinae; Bothrops.
1 (bases 1 to 461)
Junqueira-de-Azevedo, I.L.M. and Ho, P.L.
A survey of gene expression and diversity in the venom glands of
the pitviper snake Bothrops insularis through the generation of
expressed sequence tags (ESTs)
Gene 299 (1-2), 279-291 (2002)
22347338
12459276
Contact: Paulo Lee Ho
Centro de Biotecnologia
Instituto Butantan
Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
Tel: 55 11 37 26 7222 ext. 2083
Fax: 55 11 37 26 1505
Email: hople@usp.br
This EST corresponds to cluster BITL09A (see Reference)
Seq primer: M13R.

FEATURES
source
Location/Qualifiers
1. 461
/organism="Bothrops insularis"
/mol_type="mRNA"
/db_xref="taxon:8723"
/tissue type="venom glands"
/clone lib="Snake Bothrops insularis library IL2"
/note="Organ: venom glands; Vector: pGEM11zf+; Site 1: Eco
RI; Site 2: Not I; 5'ug of mRNA from Bothrops insularis
venom glands were primed with oligo-(dT) and reverse
transcribed to cDNA using Superscript Plasmid System for
cDNA Synthesis and Cloning (Life Technologies). The cDNAs
were selected by size (350-600 pb and up 600 pb) in
agarose gel electrophoresis, linked to Eco RI adapters and
directionally cloned in pGEM11zf+ vector (Promega). ESTs
were generated from random clones and grouped in unique
sequences. The putative identification of each EST or
cluster was obtained through Blast searches (e-value <
e-05)."

BASE COUNT 104 a 125 c 107 g 120 t 5 others
ORIGIN

Alignment Scores:
Pred. No.: 461
Score: 312.50
Percent Similarity: 74.47%
Best Local Similarity: 60.64%
Query Match: 43.71%
DB: 12
Length: 461
Matches: 57
Conservative: 13
Mismatch: 23
Indels: 1
Gaps: 1

US-09-938-114-2 (1-129) x BM401683 (1-461)

QY 35 AenGlyGlyHisLeuValSerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeu 54
DB 459 AATGGCGCCATCTGTTTCTATCGAAGCGCGGAGAGCAGCCCTTTGGCCAGCTG 400
QY 55 lleAlaGlnLysIleLysSerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsn 74
DB 399 GTCCNTGAGAACAAACACACAGTCGCGAATGATGTCTGATCGAATGAGGGTTCAAGGC 340
QY 75 LysGluLysGlnCysSerIleGluTrpSerAspGlySerSerIleSerLysGlnAsnTrp 94
DB 339 AAAGAAAGCAATGACGC---TCGTGGAGCGATGCTCCAGCTCAGTTATGAGAACTGG 283

Percent Similarity: 70.87% Conservative: 17
 Best Local Similarity: 54.37% Mismatches: 24
 Query Match: 42.03% Indels: 6
 DB: 12 Gaps: 1

US-09-938-114-2 (1-129) x BM401659 (1-456)

Qy 1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
 Db 149 GAGTGTCTCTGATGTCACCATANACAGTATTGCTCAAAAGTTCTTCCAAACAAAG 208
 Qy 21 LysThrTrpThr-----AspAlaGluSerPheCysThrLysGlnValAsn 35
 Db 209 GAGAGTGGGATGATGTGAGTGTATGATGAGAGGTTCTGCTCGAGCAGCGGAG 268
 Qy 36 GlyGlyHisLeuValSerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIle 55
 Db 269 GCGGCCATCTGGTCTCTATCAAGCGATGAAGAAGCAGACTTTGTGGCCCACTGGTC 328
 Qy 56 AlaGlnLysIleLysSerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLys 75
 Db 329 GCTCCGAACATANGAATCCAAATACATATGCTGGATCGGACTCGGAGATTGAAACAA 388
 Qy 76 GluLysGlnCysSerIleGluTrpSerAspGlySerSerIleSerLysGluAsn-TipI1 95
 Db 389 AAACAGCAATGACGCTCGAAGTGGAGCGATTACTCCAGGTCAGTTATGAGAACCTGGTT 448
 Qy 95 eGluGlu 97
 Db 449 CGAGGAA 455

RESULT 15

BM401401

LOCUS

DEFINITION GH025F Snake Bothrops insularis library IL2 Bothrops insularis cDNA
 5' similar to Snake venom C-type lectin, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

EST.

Bothrops insularis (island jararaca)

Bothrops insularis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;

Viperidae; Crotalinae; Bothrops.

1 (bases 1 to 654)

Junqueira-de-Azevedo, I.L.M. and Ho, P.L.

A survey of gene expression and diversity in the venom glands of

the pitviper snake Bothrops insularis through the generation of

expressed sequence tags (ESTs)

Gene 299 (1-2), 279-291 (2002)

22347338

12459276

Contact: Paulo Lee Ho

Centro de Biotecnologia

Instituto Butantan

Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900

Tel: 55 11 37 26 7222 ext. 2083

Fax: 55 11 37 26 1505

Email: hoplee@usp.br

This EST corresponds to cluster BITL10A (see Reference)

Seq primer: M13F.

Location/Qualifiers

1..654

/organism="Bothrops insularis"

/mol_type="mRNA"

/db_xref="taxon:8723"

/tissue type="venom glands"

/clone_lib="Snake Bothrops insularis library IL2"

/note="Organ: venom glands; Vector: pGEM11zf+; Site: Eco

RI; Site: 2: Not I; Sug of mRNA from Bothrops insularis

venom glands were primed with oligo-(dT) and reverse

transcribed to cDNA using Superscript plasmid system for

cDNA Synthesis and Cloning (Life Technologies). The cDNAs

FEATURES

source

were selected by size (350-600 pb and up 600 pb) in
 agarose gel electrophoresis, linked to Eco RI adapters and
 directionally cloned in pGEM11zf+ vector (Promega). ESTs
 were generated from random clones and grouped in unique
 sequences. The putative identification of each EST or
 cluster was obtained through Blast searches (e-value <
 e-05)."

BASE COUNT 154 a 156 c 175 g 161 t 8 others

ORIGIN

Alignment Scores:

Pred. No.: 5.65e-22 Length: 654

Score: 279.00 Matches: 61

Percent Similarity: 60.77% Conservative: 18

Best Local Similarity: 46.92% Mismatches: 43

Query Match: 39.02% Indels: 8

DB: 12 Gaps: 5

US-09-938-114-2 (1-129) x BM401401 (1-654)

Qy 1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
 Db 143 GATTGTCCCTGATGTGCTCTCTATGAGGAGTGTCTACAGGTTCTTCAACAAAG 202
 Qy 21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
 Db 203 ATGAATCGGAGGATGCAGAGAAATTCGCACACACAGCAGACAGCGGACATCTGGTC 262
 Qy 41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
 Db 263 TCCTTCCAGACGCGAAGACAGATTTCGTGTGTCCTCACCTCACCAATTTTGAGA 322
 Qy 61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
 Db 323 GACAGT-----TTTGTCTGGACCGGACTG-----ACGAGTCTGGAAAGGATGCAGG 370
 Qy 81 IleGluTrpSerAspGlySerSerIleSer---LysGluAsnTrpIleGluGluSer 99
 Db 371 TTCGAGTGGAGCGATGGCAGCGACCTCANCCTACAAGGACCACTACCAGTTTGTGTTTCA 430
 Qy 100 Lys---LysCysLeuGlyValHisIleGluTrpGlyPheHisLysTrpGluAsnPheTyr 118
 Db 431 GAATATGAGTGT-----GTCCATCAAGACAGAAAAATACAAATGGAGGATTATACCC 484
 Qy 119 CysGluGlnGlnAspPheValCysGlu 128
 Db 485 TCACAAAGTTGGAATATTCGTCTGCGAG 514

Search completed: December 8, 2003, 17:24:30

Job time : 3345.94 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:45:44 ; Search time 17.4 Seconds
(without alignments)
264.544 Million cell updates/sec

Title: US-09-938-114-3

Perfect score: 170
Sequence: 1 DCSWSSYEGHCYKVFQSKWTDAESF 29

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 19Jun03:*

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	170	100.0	129	ABU08798	Deinagkistrodon ac
2	138	81.2	116	AA24427	Sequence of the pl
3	135	79.4	30	AA72232	Vipera palestinae
4	135	79.4	43	AA72234	Vipera palestinae
5	135	79.4	127	AA72236	Pigmy rattlesnake
6	134	78.8	151	AA20178	N-terminal sequenc
7	132	77.6	50	AA23889	Deinagkistrodon ac
8	132	77.6	146	ABU08799	Sequence of polype
9	131	77.1	123	AA38225	

10	130	76.5	32	16	AA72231	Vipera palestinae
11	130	76.5	52	16	AA72233	Vipera palestinae
12	130	76.5	132	16	AA72235	Vipera palestinae
13	128	75.3	120	23	AB83143	Ahyliantinfarctas
14	127	74.7	30	22	AA51540	Snake venom blood
15	127	74.7	146	22	AA51544	Snake venom blood
16	127	74.7	152	23	AA520179	Pigmy rattlesnake
17	124	72.9	122	23	AB83142	Ahyliantinfarctas
18	123.5	72.6	28	13	AA23886	N-terminal sequenc
19	120	70.6	131	14	AA38222	Sequence of polype
20	119	70.0	30	22	AA51541	Snake venom blood
21	119	70.0	125	14	AA345157	Botrocetin beta su
22	119	70.0	145	23	AA014521	Korean adder snake
23	119	70.0	152	22	AA51543	Snake venom blood
24	117	68.8	133	14	AA45156	Botrocetin alpha s
25	116	68.2	30	17	AAW04861	N-terminal peptide
26	116	68.2	30	17	AAW04862	N-terminal peptide
27	116	68.2	107	14	AA38227	Sequence of polype
28	114	67.1	122	14	AA38231	Sequence of polype
29	112	65.9	132	14	AA38226	Sequence of polype
30	111	65.3	36	13	AA23885	N-terminal sequenc
31	111	65.3	38	16	AA71978	Snake venom derive
32	111	65.3	110	23	AAO20975	110-mer central lo
33	111	65.3	126	16	AA71979	Snake venom derive
34	111	65.3	126	21	AAV85627	Snake venom derive
35	111	65.3	126	23	AAO20973	126-mer wild-type
36	111	65.3	126	23	AAO20981	M8-G4 D54A mutant
37	111	65.3	126	23	AAO20982	M9-G4 Y58A mutant
38	111	65.3	126	23	AAO20983	M10-G4 K61A mutant
39	111	65.3	126	23	AAO20984	M11-G4 E62A mutant
40	111	65.3	126	23	AAO20985	M12-G4 Y63A mutant
41	111	65.3	126	23	AAO20986	M13-G4 R66A mutant
42	111	65.3	126	23	AAO20987	M14-G4 Y67A mutant
43	111	65.3	126	23	AAO20988	M15-G4 R100A mutant
44	111	65.3	126	23	AAO20989	M16-G4 D101A mutan
45	111	65.3	126	23	AAO20990	M17-G4 R103A mutan

ALIGNMENTS

RESULT 1

ABU08798

ID ABU08798 standard; Protein; 129 AA.

XX ABU08798;

XX AC

XX 02-JUN-2003 (first entry)

DT

XX Deinagkistrodon acutus antithrombosis enzyme alpha chain.

DE

XX

XX

KW

KW

KW

KW

KW

KW

KW

XX

OS

XX

XX

FT

FT

FT

FT

FT

FT

XX

XX

PN

XX

Location/Qualifiers
1..9
/note= "Antithrombosis enzyme alpha chain N-terminus not encoded by abx93668"
1..29
/note= "Specifically Claimed in claim 7"
91
/note= "Encoded by TAT"

US6489451-B1.

PD 03-DEC-2002.
 XX
 XX
 PF 10-APR-1998; 98US-0058740.
 XX
 XX
 PR 10-APR-1997; 97US-043886P.
 XX
 XX
 PA (HBEF-) HBEFI SIU FUNG USTC PHARM CO LTD.
 XX
 XX
 PI Li BX, Cheng X;
 XX
 XX WPI; 2003-352116/33.
 DR N-PSDB; ABX93668.
 XX
 XX
 PT New purified Agkistrodon actus anti-thrombosis enzyme, useful for
 PT preventing and treating vaso-occlusive and thromboembolic disorders,
 PT including myocardial infarction, restenosis, cerebral thrombosis and
 PT unstable angina -
 XX
 PS Claim 12; Column 11; 19pp; English.
 XX
 CC The invention relates to a new Deinagkistrodon actus anti-thrombosis
 CC enzyme, where the enzyme hydrolyses fibrin, dissolves blood clots and
 CC prevents platelet aggregation. The anti-thrombosis enzyme was
 CC administered to rabbits intravenously. Thrombosis was determined before
 CC and after the administration at 0.5, 1.0, 2.0 and 3.0 hours by the
 CC Chandler method. The enzyme showed anti-thrombosis activity at 0.5 hour
 CC following administration at 0.005 micro/kg and this activity was
 CC increased significantly at 1.0 hour and at 0.01 micro /kg. The
 CC composition is useful for preventing and treating vaso-occlusive and
 CC thromboembolic disorders, including myocardial infarction, restenosis,
 CC angiotatic thrombosis, cerebral thrombosis, ischaemic cerebral vascular
 CC diseases, unstable angina, acute thrombosis, unstable stenocardia,
 CC thromboangiitis obliterans, pulmonary embolism, deep vein thrombosis,
 CC peripheral arterial occlusion, stroke. It is also useful for treating
 CC atherosclerosis, oedema and inflammation, cancer and neurodegenerative
 CC diseases. The present sequence represents the amino acid sequence of the
 CC Deinagkistrodon actus antithrombosis enzyme alpha chain.
 XX
 SQ Sequence 129 AA;
 Query Match 100.0%; Score 170; DB 24; Length 129;
 Best Local Similarity 100.0%; Pred. No. 8.3e-16;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DCSDDWSYEGHCYKVKQSKTWTDAESF 29
 DB 1 DCSDDWSYEGHCYKVKQSKTWTDAESF 29
 RESULT 2
 AAR24427
 ID AAR24427 standard; Protein; 116 AA.
 XX
 XX
 AC AAR24427;
 XX
 XX 25-MAR-2003 (updated)
 DT 21-NOV-1992 (first entry)
 XX
 XX Sequence of the platelet glycoprotein GPIb inhibitor beta chain
 (CHH-B-beta).
 DE
 DE Platelet adherence inhibitor; platelet anti-adhesive;
 XX antithrombotic agent; von Willebrand Factor;
 KW platelet glycoprotein GPIb-IX complex.
 KW
 XX Crotalus horridus horridus.
 XX
 XX W09208472-A1.
 PN
 XX 29-MAY-1992.
 PD
 XX 14-NOV-1991; 91WO-US08516.
 PF
 XX

PR 16-NOV-1990; 90US-0614443.
 XX
 XX (CORT-) COR THERAPEUTICS INC.
 XX
 XX Scarbrough RM;
 XX
 DR WPI; 1992-199936/24.
 XX
 XX Platelet antiadhesive peptide(s) obtd. from snake venom - also
 PT inhibit thrombus formation; for treatment of arteriosclerosis,
 PT atherosclerosis, acute myocardial infarction, chronic unstable
 PT angina, etc.
 XX
 PS Example; Fig 6; 5pp; English.
 XX
 CC The PAA was purified from a solution of snake venom. Analysis of
 CC the peak inhibitory fractions by SDS-PAGE revealed 2 major proteins
 CC migrating with a mol. wt. 23-28 kD. Both were able to inhibit
 CC botrocetin and ristocetin induced platelet agglutination. They were
 CC called CHH-A and CHH-B. The earlier eluting subunit (CHH-B-beta)
 CC and later eluting subunit (CHH-B-alpha) were individually submitted
 CC to N-terminal sequence analysis. The complete amino acid sequence
 CC for the alpha and beta chains are given in AAR24426 and AAR24427.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 116 AA;
 Query Match 81.2%; Score 138; DB 13; Length 116;
 Best Local Similarity 75.9%; Pred. No. 2e-11;
 Matches 22; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 QY 1 DCSDDWSYEGHCYKVKQSKTWTDAESF 29
 DB 1 DCSDDWSYEGHCYKVKQSKTWTDAESF 29
 RESULT 3
 AAR72232
 ID AAR72232 standard; peptide; 30 AA.
 XX
 XX
 AC AAR72232;
 XX
 XX 25-MAR-2003 (updated)
 DT 04-DEC-1995 (first entry)
 XX
 XX Vipera palestinae venom derived peptide N-terminal fragment.
 XX Vipera palestinae; venom; N-terminal fragment; antithrombotic agent;
 KW platelet binding; von Willebrand factor.
 XX
 XX Vipera palestinae.
 OS
 XX W09509183-A1.
 PN
 XX 06-APR-1995.
 PD
 XX 27-SEP-1994; 94WO-JP01583.
 PF
 XX 28-SEP-1993; 93JP-0241666.
 PR 14-JAN-1994; 94JP-0002691.
 PR 10-JUN-1994; 94JP-0128518.
 XX
 XX (AJIN) AJINOMOTO KK.
 PA
 XX Fukuchi N, Ishii K, Kaide K, Kobayashi T;
 PI WPI; 1995-147392/19.
 DR
 XX Peptide derived from Vipera palestinae venom - inhibits binding
 PT of platelets to von Willebrand factor, useful as antithrombotic
 PT agent
 XX
 PS Claim 3; Page 17; 37pp; Japanese.

```

XX AAR72232 is a Vipera palestinae venom derived N-terminal fragment,
CC the full peptide is given in AAR72236. A peptide compsn. comprising
CC AAR72236 and AAR72235 is claimed, it inhibits the binding of platelets
CC to von Willebrand factors and can therefore be used as an anti-
CC thrombotic agent.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX SQ Sequence 30 AA;
Query Match 79.4%; Score 135; DB 16; Length 30;
Best Local Similarity 75.9%; Pred. No. 1.1e-11;
Matches 22; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 DCSDDSSSYEGHCYKVKFKQSKTWTDAESF 29
Db 1 DCPDSSSHGHCYKVFNLDTWADAEKF 29

RESULT 4
AAR72234
ID AAR72234 standard; peptide; 43 AA.
XX
XX AC AAR72234;
XX
XX DT 25-MAR-2003 (updated)
XX DT 04-DEC-1995 (first entry)
XX
XX Vipera palestinae venom derived peptide N-terminal fragment.
XX
XX Vipera palestinae; venom; N-terminal fragment; antithrombotic agent;
XX platelet binding; von Willebrand factor.
XX
XX Vipera palestinae.
XX
XX WO9509183-A1.
XX
XX PD 06-APR-1995.
XX
XX PF 27-SEP-1994; 94WO-JP01583.
XX
XX PR 28-SEP-1993; 93JP-0241666.
XX PR 14-JAN-1994; 94JP-0002691.
XX PR 10-JUN-1994; 94JP-0128518.
XX
XX (AJIN ) AJINOMOTO KK.
XX
XX PI Fukuchi N, Ishii K, Kaيدا K, Kobayashi T;
XX
XX WPI; 1995-147392/19.
XX
XX Peptide derived from Vipera palestinae venom - inhibits binding
XX of platelets to von Willebrand factor, useful as antithrombotic
XX agent
XX
XX Claim 4; Page 18; 37pp; Japanese.
XX
XX AAR72234 is a Vipera palestinae venom derived N-terminal fragment,
XX the full peptide is given in AAR72236. A peptide compsn. comprising
XX AAR72236 and AAR72235 is claimed, it inhibits the binding of platelets
XX to von Willebrand factors and can therefore be used as an anti-
XX thrombotic agent.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX SQ Sequence 43 AA;
Query Match 79.4%; Score 135; DB 16; Length 43;
Best Local Similarity 75.9%; Pred. No. 1.1e-11;
Matches 22; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 DCSDDSSSYEGHCYKVKFKQSKTWTDAESF 29
Db 1 DCPDSSSHGHCYKVFNLDTWADAEKF 29

```

```

RESULT 5
AAR72236
ID AAR72236 standard; peptide; 127 AA.
XX
XX AC AAR72236;
XX
XX DT 25-MAR-2003 (updated)
XX DT 04-DEC-1995 (first entry)
XX
XX Vipera palestinae venom derived peptide.
XX
XX Vipera palestinae; venom; antithrombotic agent;
XX platelet binding; von Willebrand factor.
XX
XX OS Vipera palestinae.
XX
XX WO9509183-A1.
XX
XX PD 06-APR-1995.
XX
XX PF 27-SEP-1994; 94WO-JP01583.
XX
XX PR 28-SEP-1993; 93JP-0241666.
XX PR 14-JAN-1994; 94JP-0002691.
XX PR 10-JUN-1994; 94JP-0128518.
XX
XX (AJIN ) AJINOMOTO KK.
XX
XX PI Fukuchi N, Ishii K, Kaيدا K, Kobayashi T;
XX
XX WPI; 1995-147392/19.
XX
XX Peptide derived from Vipera palestinae venom - inhibits binding
XX of platelets to von Willebrand factor, useful as antithrombotic
XX agent
XX
XX Claim 5; Pages 19-20; 37pp; Japanese.
XX
XX AAR72236 is a Vipera palestinae venom derived peptide. A peptide
XX compsn. comprising AAR72235 and AAR72236 is claimed, it inhibits the
XX binding of platelets to von Willebrand factors and can therefore
XX be used as an anti- thrombotic agent.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX SQ Sequence 127 AA;
Query Match 79.4%; Score 135; DB 16; Length 127;
Best Local Similarity 75.9%; Pred. No. 5.6e-11;
Matches 22; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 DCSDDSSSYEGHCYKVKFKQSKTWTDAESF 29
Db 1 DCPDSSSHGHCYKVFNLDTWADAEKF 29

RESULT 6
AAE20178
ID AAE20178 standard; Protein; 151 AA.
XX
XX AC AAE20178;
XX
XX DT 18-JUN-2002 (first entry)
XX
XX DE Pigmy rattlesnake venom gland protein, Zsnk2.
XX
XX Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;
XX platelet aggregation; Zsnk2.
XX
XX OS Sistrurus miliarius.
XX
XX PH Key Location/Qualifiers

```

FT Peptide 1..19
 FT /label= Signal_peptide
 FT Protein 20..151
 FT /label= Mature_Zsnk2_protein
 FT Disulfide-bond 21..32
 FT Disulfide-bond 49..140
 FT Disulfide-bond 115..132
 XX
 PN W0200214364-A2.
 XX
 XX 21-FEB-2002.
 XX
 XX 13-AUG-2001; 2001WO-US25310.
 XX
 XX 14-AUG-2000; 2000US-225072P.
 PR 14-AUG-2000; 2000US-225087P.
 PR 15-AUG-2000; 2000US-225489P.
 PR 15-AUG-2000; 2000US-225490P.
 PR 20-DEC-2000; 2000US-356997P.
 XX
 XX (ZYMO) ZYMOGENETICS INC.
 XX
 XX Sheppard PO, Bishop PD;
 PI
 XX WPI; 2002-269180/31.
 XX
 XX N-PSDB; AAD32053.
 XX
 XX New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which
 FT affect blood coagulation and platelet aggregation system, useful in
 FT therapy and diagnostics, or as tools in the study of genetics or
 FT molecular biology -
 XX
 PS Claim 3; Page 2; 79pp; English.
 XX
 CC The invention relates to new pigmy rattlesnake (Sistrurus miliarius)
 CC venom gland proteins, which affect blood coagulation and platelet
 CC aggregation system. The polypeptides, which affect blood coagulation and
 CC platelet aggregation system, are useful in therapy and diagnostics. The
 CC polypeptides are also useful as an educational tool in laboratory
 CC practical kits for courses related to genetics and molecular biology,
 CC protein chemistry and antibody production and analysis. The
 CC polynucleotide or polypeptide can be used as standards or as unknowns
 CC for testing purposes. The polypeptides are also useful in identifying
 CC proteins by western blotting, protein purification, determining the
 CC weight of expressed polypeptides as a ratio to total protein expressed,
 CC identifying peptide cleavage sites, coupling amino and carboxyl terminal
 CC tags, mass spectrometry, circular dichroism to determine conformation or
 CC affinity chromatography columns to purify the protein, cloning or
 CC sequencing. The present sequence is Sistrurus miliarius venom gland
 CC protein, Zsnk2.
 XX
 XX Sequence 151 AA;
 XX
 Query Match 78.8%; Score 134; DB 23; Length 151;
 Best Local Similarity 75.9%; Pred. No. 9.5e-11;
 Matches 22; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 QY 1 DCSSDWSYEGHCYKVFQSKTWTDAESF 29
 Db 20 DCPDWSYDQHCYKVFSELKTWDDAESF 48
 XX
 RESULT 7
 AAR23889
 ID AAR23889 standard; Protein; 50 AA.
 XX
 AC AAR23889;
 XX
 XX 25-MAR-2003 (updated)
 DT 21-NOV-1992 (first entry)
 XX
 XX N-terminal sequence of the earlier eluting subunit PP-beta of
 DE the later eluting peak of the platelet antiadhesive peptide (PAA).

XX Platelet adherence inhibitor; platelet anti-adhesive;
 KW antithrombotic agent; von Willebrand Factor;
 KW platelet glycoprotein GPIb-IX complex.
 XX
 OS Pseudocercariae persicus.
 XX
 PN W09208472-A1.
 XX
 XX 29-MAY-1992.
 XX
 XX 14-NOV-1991; 91WO-US08516.
 XX
 XX 16-NOV-1990; 90US-0614443.
 PR (CORT-) COR THERAPEUTICS INC.
 XX
 XX Scarborough RM;
 PI
 XX WPI; 1992-199936/24.
 DR
 XX Platelet antiadhesive peptide(s) obt'd. from snake venom - also
 FT inhibit thrombus formation; for treatment of arteriosclerosis,
 FT atherosclerosis, acute myocardial infarction, chronic unstable
 FT angina, etc.
 XX
 PS Claim 7; Page 30; 5pp; English.
 XX
 CC The PAA was purified from a solution of snake venom. The earlier
 CC eluting subunit, PP-beta was subjected to Edman degradation for 50
 CC cycles to obtain the sequence in AAR23889. The later eluting
 CC sequence - the PP-alpha chain was determined for 31 cycles to give
 CC the SQ in AAR23890.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 XX Sequence 50 AA;
 XX
 Query Match 77.6%; Score 132; DB 13; Length 50;
 Best Local Similarity 75.9%; Pred. No. 5e-11; Indels 0; Gaps 0;
 Matches 22; Conservative 1; Mismatches 6;
 QY 1 DCSSDWSYEGHCYKVFQSKTWTDAESF 29
 Db 1 DCPDWSYDQHCYKVFSELKTWDDAESF 29
 XX
 RESULT 8
 ABU08799
 ID ABU08799 standard; Protein; 146 AA.
 XX
 AC ABU08799;
 XX
 XX 02-JUN-2003 (first entry)
 DT
 XX Deinagkistrodon acutus antithrombosis enzyme beta chain.
 DE
 XX
 KW Antithrombosis; beta chain; fibrin hydrolysis; blood clot; enzyme;
 KW platelet aggregation; vaso-occlusive disorder; thromboembolic disorder;
 KW myocardial infarction; restenosis; cancer; neurodegenerative disease;
 KW angioathic thrombosis; cerebral thrombosis; thromboangitis obliterans;
 KW ischaemic cerebral vascular disease; unstable angina; acute thrombosis;
 KW unstable stenocardia; pulmonary embolism; deep vein thrombosis; oedema;
 KW peripheral arterial occlusion; stroke; atherosclerosis; inflammation;
 KW thrombosis.
 XX
 XX Deinagkistrodon acutus.
 OS
 XX
 Key Location/Qualifiers
 FH 1..23
 FT Peptide /note= "Leader peptide"
 FT Protein 24..146
 FT /note= "Mature antithrombosis enzyme beta chain. Residues
 FT 24 to 40 specifically claimed in claim 8"

FT	Misc-difference	72	
FT	/label= Unknown		
FT	/note= "Encoded by TNN"		
FT	Misc-difference	72	
FT	/label= Unknown		
FT	/note= "Encoded by NNN"		
FT	Misc-difference	72	
FT	/label= Unknown		
FT	/note= "Encoded by NNT"		
XX			
PN	US6489451-B1.		
XX			
PD	03-DEC-2002.		
XX			
PP	10-APR-1998;	98US-0058740.	
XX			
PR	10-APR-1997;	97US-043886P.	
XX			
PA	(HEFE-) HEFEI SIU FUNG USTC PHARM CO LTD.		
XX			
PI	Li BX, Cheng X;		
XX			
DR	WPI; 2003-352116/33.		
DR	N-PSDB; ABX93674.		
XX			
PT	New purified Aqkistrodon actus anti-thrombosis enzyme, useful for		
PT	preventing and treating vaso-occlusive and thromboembolic disorders,		
PT	including myocardial infarction, restenosis, cerebral thrombosis and		
PT	unstable angina		
XX			
PS	Disclosure; Fig 1; 19pp; English.		
XX			
CC	The invention relates to a new Deinagkistrodon actus anti-thrombosis		
CC	enzyme, where the enzyme hydrolyses fibrin, dissolves blood clots and		
CC	prevents platelet aggregation. The anti-thrombosis enzyme was		
CC	administered to rabbits intravenously. Thrombosis was determined before		
CC	and after the administration at 0.5, 1.0, 2.0 and 3.0 hours by the		
CC	Chandler method. The enzyme showed anti-thrombosis activity at 0.5 hour		
CC	following administration at 0.005 micro/kg and this activity was		
CC	increased significantly at 1.0 hour and at 0.01 micro /kg. The		
CC	composition is useful for preventing and treating vaso-occlusive and		
CC	thromboembolic disorders, including myocardial infarction, restenosis,		
CC	angiopathic thrombosis, cerebral thrombosis, ischaemic cerebral vascular		
CC	diseases, unstable angina, acute thrombosis, unstable stenocardia,		
CC	thromboangiitis obliterans, pulmonary embolism, deep vein thrombosis,		
CC	peripheral arterial occlusion, stroke. It is also useful for treating		
CC	atherosclerosis, oedema and inflammation, cancer and neurodegenerative		
CC	diseases. The present sequence represents the amino acid sequence of the		
CC	Deinagkistrodon actus antithrombosis enzyme beta chain.		
XX			
SQ	Sequence	146 AA;	
	Query Match	77.6%; Score 132; DB 24; Length 146;	
	Best Local Similarity	72.4%; Pred. No. 1.7e-10;	
	Matches	21; Conservative	2; Mismatches 6; Indels 0; Gaps 0;
QY	1	DCSSDWSSYEGHCYKVPKOSKTWDAESF	29
		: :	
Db	24	DCPSEWSSYEGHCYKFPDKPTWADAKEF	52
		: :	
RESULT	9		
AA	AAAR38225		
XX	AAAR38225 standard; protein; 123 AA.		
XX	AAAR38225;		
XX			
DT	25-MAR-2003 (updated)		
DT	01-OCT-1993 (first entry)		
XX			

PA (AJIN) AJINOMOTO KK.
 XX PI Fukuchi N, Ishii K, Kaide K, Kobayashi T;
 XX WPI; 1995-147392/19.
 XX Query Match 76.5%; Score 130; DB 16; Length 32;
 XX Best Local Similarity 72.4%; Pred. No. 5.7e-11;
 XX Matches 21; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 XX Peptide derived from Vipera palestinae venom - inhibits binding
 PT of platelets to von Willebrand factor, useful as antithrombotic
 PT agent
 XX Claim 3; Page 17; 37pp; Japanese.
 XX AAR72231 is a Vipera palestinae venom derived N-terminal fragment.
 CC the full peptide is given in AAR72235. A peptide compsn. comprising
 CC AAR72235 and AAR72236 is claimed, it inhibits the binding of platelets
 CC to von Willebrand factors and can therefore be used as an anti-
 CC thrombotic agent.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX SQ Sequence 32 AA;
 XX Query Match 76.5%; Score 130; DB 16; Length 32;
 XX Best Local Similarity 72.4%; Pred. No. 5.7e-11;
 XX Matches 21; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 XX 1 DCSDDSSSYEGHCYKVFQKSKTWTDAESF 29
 XX |||||:|||||: |||: |||
 Db 3 DCPDSSSHGHCYKVFRLFKTWEAEKF 31
 RESULT 11
 AAR72233
 ID AAR72233 standard; peptide; 52 AA.
 XX AC AAR72233;
 XX 25-MAR-2003 (updated)
 DT 04-DEC-1995 (first entry)
 XX Vipera palestinae venom derived peptide N-terminal fragment.
 DE Vipera palestinae; venom; N-terminal fragment; antithrombotic agent;
 XX platelet binding; von Willebrand factor.
 XX Vipera palestinae.
 OS WO9509183-A1.
 XX 06-APR-1995.
 XX 27-SEP-1994; 94WO-JP01583.
 XX 28-SEP-1993; 93JP-0241666.
 PR 14-JAN-1994; 94JP-0002691.
 PR 10-JUN-1994; 94JP-0128518.
 XX (AJIN) AJINOMOTO KK.
 XX PI Fukuchi N, Ishii K, Kaide K, Kobayashi T;
 XX WPI; 1995-147392/19.
 XX Peptide derived from Vipera palestinae venom - inhibits binding
 PT of platelets to von Willebrand factor, useful as antithrombotic
 PT agent
 XX Claim 4; Page 18; 37pp; Japanese.
 XX AAR72233 is a Vipera palestinae venom derived N-terminal fragment.
 CC the full peptide is given in AAR72235. A peptide compsn. comprising
 CC AAR72235 and AAR72236 is claimed, it inhibits the binding of platelets
 CC to von Willebrand factors and can therefore be used as an anti-
 CC thrombotic agent.

CC (Updated on 25-MAR-2003 to correct PN field.)
 XX SQ Sequence 52 AA;
 XX Query Match 76.5%; Score 130; DB 16; Length 52;
 XX Best Local Similarity 72.4%; Pred. No. 9.9e-11;
 XX Matches 21; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 XX 1 DCSDDSSSYEGHCYKVFQKSKTWTDAESF 29
 XX |||||:|||||: |||: |||
 Db 3 DCPDSSSHGHCYKVFRLFKTWEAEKF 31
 RESULT 12
 AAR72235
 ID AAR72235 standard; peptide; 132 AA.
 XX AC AAR72235;
 XX 25-MAR-2003 (updated)
 DT 04-DEC-1995 (first entry)
 XX Vipera palestinae venom derived peptide.
 DE Vipera palestinae; venom; antithrombotic agent;
 XX platelet binding; von Willebrand factor.
 XX Vipera palestinae.
 OS WO9509183-A1.
 XX 06-APR-1995.
 XX 27-SEP-1994; 94WO-JP01583.
 XX 28-SEP-1993; 93JP-0241666.
 PR 14-JAN-1994; 94JP-0002691.
 PR 10-JUN-1994; 94JP-0128518.
 XX (AJIN) AJINOMOTO KK.
 XX PI Fukuchi N, Ishii K, Kaide K, Kobayashi T;
 XX WPI; 1995-147392/19.
 XX Peptide derived from Vipera palestinae venom - inhibits binding
 PT of platelets to von Willebrand factor, useful as antithrombotic
 PT agent
 XX Claim 5; Page 19; 37pp; Japanese.
 XX AAR72235 is a Vipera palestinae venom derived peptide. A peptide
 CC compsn. comprising AAR72235 and AAR72236 is claimed, it inhibits the
 CC binding of platelets to von Willebrand factors and can therefore
 CC be used as an anti-thrombotic agent.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX SQ Sequence 132 AA;
 XX Query Match 76.5%; Score 130; DB 16; Length 132;
 XX Best Local Similarity 72.4%; Pred. No. 2.9e-10;
 XX Matches 21; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 XX 1 DCSDDSSSYEGHCYKVFQKSKTWTDAESF 29
 XX |||||:|||||: |||: |||
 Db 3 DCPDSSSHGHCYKVFRLFKTWEAEKF 31
 RESULT 13
 ABB83143
 ID ABB83143 standard; protein; 120 AA.
 XX AC ABB83143;

Search completed: December 8, 2003, 09:52:34
Job time : 17.4 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:50:40 ; Search time 6.29714 Seconds
(without alignments)
194.853 Million cell updates/sec

Title: US-09-938-114-3

Perfect score: 170

Sequence: 1 DCSSDSSYRGHCYKVKQKWTDAESF 29

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	170	100.0	29	4	US-09-058-740-3
2	170	100.0	129	4	US-09-058-740-2
3	138	81.2	117	1	US-07-614-443A-2
4	138	81.2	117	1	US-08-294-859-2
5	138	81.2	117	1	US-08-481-676-2
6	132	77.6	50	1	US-07-614-443A-5
7	132	77.6	50	1	US-08-294-859-5
8	132	77.6	50	1	US-08-481-676-5
9	131	77.1	123	1	US-07-893-929A-4
10	131	77.1	123	5	PCT-US92-10344-4
11	120	70.6	131	1	US-07-893-929A-1
12	120	70.6	131	5	PCT-US92-10344-1
13	116	68.2	30	2	US-08-894-403-3
14	116	68.2	30	2	US-08-894-403-4
15	116	68.2	107	1	US-07-893-929A-6
16	116	68.2	107	5	PCT-US92-10344-6
17	114	67.1	123	1	US-07-893-929A-10
18	114	67.1	123	5	PCT-US92-10344-10
19	113	66.5	128	1	US-07-893-929A-8
20	113	66.5	128	5	PCT-US92-10344-8
21	112	65.9	132	1	US-07-893-929A-5
22	112	65.9	132	5	PCT-US92-10344-5
23	111	65.3	38	2	US-08-612-840A-1
24	111	65.3	126	2	US-08-612-840A-2
25	111	65.3	127	1	US-07-614-443A-1
26	111	65.3	127	1	US-08-294-859-1
27	111	65.3	127	1	US-08-481-676-1

28 111 65.3 130 1 US-07-893-929A-7 Sequence 7, Appli
29 111 65.3 130 5 PCT-US92-10344-7 Sequence 7, Appli
30 111 65.3 149 2 US-08-612-840A-8 Sequence 8, Appli
31 107 62.9 125 1 US-07-893-929A-3 Sequence 3, Appli
32 107 62.9 125 5 PCT-US92-10344-3 Sequence 3, Appli
33 104 61.2 31 1 US-07-614-443A-6 Sequence 6, Appli
34 104 61.2 31 1 US-08-294-859-6 Sequence 6, Appli
35 104 61.2 31 1 US-08-481-676-6 Sequence 6, Appli
36 103 60.6 21 2 US-08-612-840A-3 Sequence 3, Appli
37 102.5 60.3 133 1 US-07-893-929A-9 Sequence 9, Appli
38 102.5 60.3 133 5 PCT-US92-10344-9 Sequence 9, Appli
39 94 55.3 134 1 US-07-893-929A-2 Sequence 2, Appli
40 94 55.3 134 5 PCT-US92-10344-2 Sequence 2, Appli
41 84 49.4 1257 1 US-08-340-428B-49 Sequence 49, Appli
42 82 48.2 908 5 PCT-US95-03747-3 Sequence 3, Appli
43 81 47.6 17 4 US-09-058-740-4 Sequence 4, Appli
44 79 46.5 912 5 PCT-US95-03747-2 Sequence 2, Appli
45 44.1 2409 6 5180808-2 Patent No. 5180808

ALIGNMENTS

RESULT 1
US-09-058-740-3
; Sequence 3, Application US/09058740
; Patent No. 6489451
; GENERAL INFORMATION:
; APPLICANT: Chun Wang, Benjamin XY Li, Xin Cheng,
; Jing Liu, Li-Wen Niu, Wan-Zhi Huang,
; Zhen-Yu Xu, Dan Luo, Lian-Di Kang,
; Jin-Guo Ding, Fang Rong, Yan Liu and
; Hui-Ran Chen
; TITLE OF INVENTION: AN ANTITHROMBOSIS ENZYME FROM THE SNAKE
; VENOM OF AGKISTRODON ACUTUS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/058,740
; FILING DATE: 10-Apr-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Chen, Anthony C.
; REGISTRATION NUMBER: 38,673
; REFERENCE/DOCKET NUMBER: 233/298
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-058-740-3

Query Match 100.0%; Score 170; DB 4; Length 29;
Best Local Similarity 100.0%; Pred. No. 2.3e-16;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DCSSDSSSYEGHCYKVKFQSKTWTDAESF 29
Db 1 DCSSDSSSYEGHCYKVKFQSKTWTDAESF 29

RESULT 2

US-09-058-740-2
; Sequence 2, Application US/09058740
; Patent No. 6489451
; GENERAL INFORMATION:
; APPLICANT: Chun Wang, Benjamin XY Li, Xin Cheng,
; Jing Liu, Li-Wen Niu, Wan-Zhi Huang,
; Zhen-Yu Xu, Dan Luo, Lian-Di Kang,
; Jin-Guo Ding, Fang Rong, Yan Liu and
; Hui-Ran Chen
; TITLE OF INVENTION: AN ANTITHROMBOSIS ENZYME FROM THE SNAKE
; VENOM OF AGKISTRODON ACUTUS

NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSeq for Windows 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/058,740
FILING DATE: 10-Apr-1998

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:

APPLICATION NUMBER: <Unknown>

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Chen, Anthony C.

REGISTRATION NUMBER: 38,673

REFERENCE/DOCKET NUMBER: 233/298

TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440

TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 129 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: Peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-058-740-2

Query Match 100.0%; Score 170; DB 4; Length 129;
Best Local Similarity 100.0%; Pred. No. 1.1e-15;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DCSSDSSSYEGHCYKVKFQSKTWTDAESF 29
Db 1 DCSSDSSSYEGHCYKVKFQSKTWTDAESF 29

RESULT 3

US-07-614-443A-2

; Sequence 2, Application US/07614443A
; Patent No. 5342830
; GENERAL INFORMATION:
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: ANTITHROMBOSIS AGENTS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/614,443A
; FILING DATE: 19901116
; CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:
NAME: Murashige, Kate H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 22803-20003.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 117 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

US-07-614-443A-2

Query Match 81.2%; Score 138; DB 1; Length 117;
Best Local Similarity 75.9%; Pred. No. 2e-11;
Matches 22; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 DCSSDSSSYEGHCYKVKFQSKTWTDAESF 29
Db 1 DCPSDSSYEGHCYRVFQEWTDAAKXF 29

RESULT 4

US-08-294-859-2
; Sequence 2, Application US/08294859
; Patent No. 5679542
; GENERAL INFORMATION:
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: ANTITHROMBOSIS AGENTS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/294,859
; FILING DATE: 29-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.

REGISTRATION NUMBER: 29,959
 REFERENCE/DOCKET NUMBER: 2803-0003.10
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 887-1500
 TELEFAX: (202) 887-0763
 TELEX: 90-4030
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 117 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear

US-08-294-859-2

Query Match 81.2%; Score 138; DB 1; Length 117;
 Best Local Similarity 75.9%; Pred. No. 2e-11;
 Matches 22; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 DCSSDWSYEGHCYKVFQKSKTWTDAESF 29
 |||||
 Db 1 DCPDWSYEGHCYKVFQKSKTWTDAESF 29

RESULT 5

US-08-481-676-2
 Sequence 2, Application US/08481676
 Patent No. 5744584
 GENERAL INFORMATION:
 APPLICANT: SCARBOROUGH, ROBERT M.
 TITLE OF INVENTION: ANTITHROMBOSIS AGENTS
 NUMBER OF SEQUENCES: 7
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Morrison & Foerster
 STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
 CITY: Washington, D.C.
 COUNTRY: USA
 ZIP: 20006-1888

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/481,676
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 514

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/294,859
 FILING DATE: 29-AUG-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Murashige, Kate H.
 REGISTRATION NUMBER: 29,959
 REFERENCE/DOCKET NUMBER: 2803-0003.10
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 887-1500
 TELEFAX: (202) 887-0763
 TELEX: 90-4030
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 117 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear

US-08-481-676-2

Query Match 81.2%; Score 138; DB 1; Length 117;
 Best Local Similarity 75.9%; Pred. No. 2e-11;
 Matches 22; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 DCSSDWSYEGHCYKVFQKSKTWTDAESF 29
 |||||
 Db 1 DCPDWSYEGHCYKVFQKSKTWTDAESF 29

RESULT 6

US-07-614-443A-5
 Sequence 5, Application US/07614443A
 Patent No. 5342830
 GENERAL INFORMATION:
 APPLICANT: SCARBOROUGH, ROBERT M.
 TITLE OF INVENTION: ANTITHROMBOSIS AGENTS
 NUMBER OF SEQUENCES: 7
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Morrison & Foerster
 STREET: 755 Page Mill Road
 CITY: Palo Alto
 STATE: California
 COUNTRY: USA
 ZIP: 94304-1018
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/614,443A
 FILING DATE: 19901116
 CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:
 NAME: Murashige, Kate H.
 REGISTRATION NUMBER: 29,959
 REFERENCE/DOCKET NUMBER: 22803-20003.00
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 813-5600
 TELEFAX: (415) 494-0792
 TELEX: 706141
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 50 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear

US-07-614-443A-5

Query Match 77.8%; Score 132; DB 1; Length 50;
 Best Local Similarity 75.9%; Pred. No. 5e-11;
 Matches 22; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 DCSSDWSYEGHCYKVFQKSKTWTDAESF 29
 |||||
 Db 1 DCPDWSYEGHCYKVFQKSKTWTDAESF 29

RESULT 7

US-08-294-859-5
 Sequence 5, Application US/08294859
 Patent No. 5679542
 GENERAL INFORMATION:
 APPLICANT: SCARBOROUGH, ROBERT M.
 TITLE OF INVENTION: ANTITHROMBOSIS AGENTS
 NUMBER OF SEQUENCES: 7
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Morrison & Foerster
 STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
 CITY: Washington, D.C.
 COUNTRY: USA
 ZIP: 20006-1888

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/294,859
 FILING DATE: 29-AUG-1994
 CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Murashige, Kate H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 2803-0003.10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 50 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-294-859-5

Query Match 77.6%; Score 132; DB 1; Length 50;
Best Local Similarity 75.9%; Pred. No. 5e-11; Indels 0; Gaps 0;
Matches 22; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 DCSSDSSSYEGHCYKVFQKSKTWTDAESF 29
||| |||||:||||| ||| |||
Db 1 DCPDSSSHGHCYKVFNLKWTWDAEKF 29

RESULT 8
US-08-481-676-5
Sequence 5, Application US/08481676
Patent No. 5744584
GENERAL INFORMATION:
APPLICANT: SCARBOROUGH, ROBERT M.
TITLE OF INVENTION: ANTI-THROMBOSIS AGENTS
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20006-1888
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481.676
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/294,859
FILING DATE: 29-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Murashige, Kate H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 2803-0003.10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 50 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-481-676-5

Query Match 77.6%; Score 132; DB 1; Length 50;
Best Local Similarity 75.9%; Pred. No. 5e-11; Indels 0; Gaps 0;
Matches 22; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 DCSSDSSSYEGHCYKVFQKSKTWTDAESF 29
||| |||||:||||| ||| |||
Db 1 DCPDSSSHGHCYKVFNLKWTWDAEKF 29

RESULT 9
US-07-893-929A-4
Sequence 4, Application US/07893929A
Patent No. 5336667
GENERAL INFORMATION:
APPLICANT: Kirby, Edward P.
TITLE OF INVENTION: Platelet Agonists Which Bind To Platelet
TITLE OF INVENTION: Membrane Glycoprotein Ib
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Temple University - Of The Common-
wealth System of Higher Education
STREET: 406 University Services Building
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: U.S.A.
ZIP: 19122
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/893,929A
FILING DATE: 19920605
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/803,630
FILING DATE: December 3, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A.
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 6056-126 (CIP) 1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-8383
TELEFAX: (215) 568-5549
TELEX: NO. 5336667e
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 123 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
US-07-893-929A-4

Query Match 77.1%; Score 131; DB 1; Length 123;
Best Local Similarity 69.0%; Pred. No. 1.8e-10; Indels 0; Gaps 0;
Matches 20; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 DCSSDSSSYEGHCYKVFQKSKTWTDAESF 29
||| |||||:||||| ||| |||
Db 1 DCPDSSSHGHCYKVFNEPQNWDAEKF 29

RESULT 10
PCT-US92-10344-4
Sequence 4, Application PC/TUS9210344
GENERAL INFORMATION:
APPLICANT: Kirby, Edward P.
TITLE OF INVENTION: Platelet Agonists Which Bind To Platelet
TITLE OF INVENTION: Membrane Glycoprotein Ib
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Temple University - Of The Common-
wealth System of Higher Education
STREET: 406 University Services Building
CITY: Philadelphia

	Matches	137	Conserved	
Qy	1	DCSSDWSSYEGHCYKVKFKQSKTWTDAESF	29	
		: :		
Dd	1	DCPSDWSSYDQCYRVFKRLQTWDAERF	29	

CLASSIFICATION: 435
PRIOR APPLICATION NUMBER: 07/803,630
FILING DATE: December 3, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A.
REGISTRATION NUMBER: 30,480

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:54:55 ; Search time 11.9314 Seconds
(without alignments)
452.044 Million cell updates/sec

Title: US-09-938-114-3

Perfect score: 170

Sequence: 1 DCSWSSYEGHCYKVKQSKTWTDAESF 29

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications AA:*
- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
 - 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
 - 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	170	100.0	29	11	US-09-938-114-3
2	170	100.0	129	11	US-09-938-114-2
3	134	78.8	151	10	US-09-929-230-2
4	134	78.8	151	12	US-10-226-420-2
5	127	74.7	152	10	US-09-929-230-5
6	127	74.7	152	12	US-10-226-420-5
7	111	65.3	149	10	US-09-969-763-3
8	110	64.7	144	10	US-09-929-230-8
9	110	64.7	144	12	US-10-226-420-8
10	100	58.8	110	10	US-09-969-763-10
11	97	57.1	158	10	US-09-929-230-11
12	97	57.1	158	12	US-10-226-420-11
13	96	56.5	126	10	US-09-969-763-1
14	84	49.4	52	12	US-10-029-386-31252
15	84	49.4	1321	12	US-10-241-220-82

16	82	48.2	136	10	US-09-965-528-24	Sequence 24, Appl
17	82	48.2	164	9	US-09-801-438-2	Sequence 2, Appl
18	82	48.2	164	10	US-09-893-737-20	Sequence 20, Appl
19	82	48.2	883	11	US-09-759-130B-342	Sequence 342, App
20	82	48.2	883	12	US-10-188-495-72	Sequence 72, Appl
21	82	48.2	883	15	US-10-195-970-3	Sequence 3, Appl
22	82	48.2	883	15	US-10-195-970-6	Sequence 6, Appl
23	82	48.2	883	15	US-10-189-123-72	Sequence 72, Appl
24	81	47.6	17	11	US-09-938-114-4	Sequence 4, Appl
25	77	45.3	29	10	US-09-929-230-14	Sequence 14, Appl
26	77	45.3	29	12	US-10-226-420-14	Sequence 14, Appl
27	76	44.7	211	15	US-10-050-704-266	Sequence 266, App
28	76	44.7	427	15	US-10-050-704-273	Sequence 273, App
29	76	44.7	483	15	US-10-050-704-272	Sequence 272, App
30	76	44.7	911	12	US-10-237-496-52	Sequence 52, Appl
31	76	44.7	911	12	US-10-242-074-52	Sequence 52, Appl
32	76	44.7	911	12	US-10-242-574-52	Sequence 52, Appl
33	76	44.7	911	12	US-10-243-261-52	Sequence 52, Appl
34	76	44.7	911	12	US-10-243-282-52	Sequence 52, Appl
35	76	44.7	911	12	US-10-243-402-52	Sequence 52, Appl
36	76	44.7	911	12	US-10-243-431-52	Sequence 52, Appl
37	76	44.7	911	12	US-10-245-164-52	Sequence 52, Appl
38	76	44.7	911	12	US-10-244-972-52	Sequence 52, Appl
39	76	44.7	911	12	US-10-197-942-52	Sequence 52, Appl
40	76	44.7	911	12	US-10-238-196-52	Sequence 52, Appl
41	76	44.7	911	12	US-10-245-013-52	Sequence 52, Appl
42	76	44.7	911	15	US-10-245-103-52	Sequence 52, Appl
43	76	44.7	911	15	US-10-245-107-52	Sequence 52, Appl
44	76	44.7	911	15	US-10-245-143-52	Sequence 52, Appl
45	76	44.7	911	15	US-10-245-143-52	Sequence 52, Appl

ALIGNMENTS

RESULT 1

US-09-938-114-3
; Sequence 3, Application US/09938114
; Publication No. US20030022350A1
; GENERAL INFORMATION:

APPLICANT: Chun Wang, Benjamin XY Li, Xin Cheng,
Jing Liu, Li-Wen Niu, Wan-Zhi Huang,
Zhen-Yu Xu, Dan Luo, Lian-Di Kang,
Jin-Guo Ding, Fang Rong, Yan Liu and
Hui-Kan Chen
TITLE OF INVENTION: AN ANTITHROMBOSIS ENZYME FROM THE SNAKE
VENOM OF AGKISTRODON ACUTUS

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESSES:

ADDRESSEE: LYON & LYON

STREET: 633 West Fifth Street

Suite 4700

CITY: Los Angeles

STATE: California

COUNTRY: U.S.A.

ZIP: 90071-2066

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

storage

COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: FastSeq for Windows 2.0

CURRENT APPLICATION NUMBER: US/09/938,114

FILING DATE: 23-Aug-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/058,740

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Chen, Anthony C.

REGISTRATION NUMBER: 38,673

REFERENCE/DOCKET NUMBER: 233/298

```

;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-938-114-3

Query Match      100.0%; Score 170; DB 11; Length 29;
Best Local Similarity 100.0%; Pred. No. 4.1e-16;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DCSSDSSSYEGHCYKVFQSKTWTDAESF 29
   |||||
Db 1 DCSSDSSSYEGHCYKVFQSKTWTDAESF 29
   |||||

RESULT 2
US-09-938-114-2
; Sequence 2, Application US/09938114
; Publication No. US20030022350A1
; GENERAL INFORMATION:
; APPLICANT: Chun Wang, Benjamin XY Li, Xin Cheng,
;           Jing Liu, Li-Wen Niu, Wan-Zhi Huang,
;           Zhen-Yu Xu, Dan Luo, Lian-Di Kang,
;           Jin-Guo Ding, Fang Rong, Yan Liu and
;           Hui-Ran Chen
; TITLE OF INVENTION: AN ANTI-THROMBOSIS ENZYME FROM THE SNAKE
;           VENOM OF AGKISTRODON ACUTUS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
;           Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/938,114
; FILING DATE: 23-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/059,740
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Chen, Anthony C.
; REGISTRATION NUMBER: 38,673
; REFERENCE/DOCKET NUMBER: 233/298
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 129 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-938-114-2

Query Match      100.0%; Score 170; DB 11; Length 129;
Best Local Similarity 100.0%; Pred. No. 1.9e-15;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DCSSDSSSYEGHCYKVFQSKTWTDAESF 29
   |||||
Db 1 DCSSDSSSYEGHCYKVFQSKTWTDAESF 29
   |||||

RESULT 3
US-09-929-230-2
; Sequence 2, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/09/929,230
; CURRENT FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Sistrurus miliarius
US-09-929-230-2

Query Match      78.8%; Score 134; DB 10; Length 151;
Best Local Similarity 75.9%; Pred. No. 1.6e-10;
Matches 22; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 DCSSDSSSYEGHCYKVFQSKTWTDAESF 29
   |||||
Db 20 DCPDSSSYDQHCYKVFSELKTWDDAESF 48
   |||||

RESULT 4
US-10-226-420-2
; Sequence 2, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/10/226,420
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Sistrurus miliarius
US-10-226-420-2

Query Match      78.8%; Score 134; DB 12; Length 151;
Best Local Similarity 75.9%; Pred. No. 1.6e-10;
Matches 22; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 DCSSDSSSYEGHCYKVFQSKTWTDAESF 29
   |||||
Db 20 DCPDSSSYDQHCYKVFSELKTWDDAESF 48
   |||||

RESULT 5
US-09-929-230-5
; Sequence 5, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.

```

; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/09/929,230
; CURRENT FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 152
; TYPE: PRT
; ORGANISM: Sistrurus miliarius
US-09-929-230-5

Query Match 74.7%; Score 127; DB 10; Length 152;
Best Local Similarity 72.4%; Pred. No. 1.4e-09;
Matches 21; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 DCSSDWSSEYEGHCYKVKFQSKTWTDAERF 29
 | | | | | | | | | | | | | | | | | |
Db 24 DCPGWSYYDHCHCYRVFKQLKTWDDAERF 52

RESULT 6
US-10-226-420-5
; Sequence 5, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/10/226,420
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 152
; TYPE: PRT
; ORGANISM: Sistrurus miliarius
US-10-226-420-5

Query Match 74.7%; Score 127; DB 12; Length 152;
Best Local Similarity 72.4%; Pred. No. 1.4e-09;
Matches 21; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 DCSSDWSSEYEGHCYKVKFQSKTWTDAERF 29
 | | | | | | | | | | | | | | | | | |
Db 24 DCPGWSYYDHCHCYRVFKQLKTWDDAERF 52

RESULT 7
US-09-969-763-3
; Sequence 3, Application US/09969763
; Publication No. US20020198363A1
; GENERAL INFORMATION:
; APPLICANT: FUKUCHI, NAOYUKI
; APPLICANT: KITO, MORIKAZU
; APPLICANT: KAYAHARA, TAKASHI
; APPLICANT: FUTAKI, FUMIE
; APPLICANT: ISHIKAWA, KOHKI
; APPLICANT: SUZUKI, EIICHIRO
; APPLICANT: GONDOH, KEIKO
; APPLICANT: SHIMBA, NOBUHISA
; APPLICANT: YAMADA, NAOYUKI
; TITLE OF INVENTION: PROTEIN HAVING ANTITHROMBOTIC ACTIVITY AND METHOD FOR PRODUCING IT
; FILE REFERENCE: 214760SU0
; CURRENT APPLICATION NUMBER: US/09/969,763
; CURRENT FILING DATE: 2000-10-25
; PRIOR APPLICATION NUMBER: JP 2000-305279
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3

```

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES
;
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
;
; FILE REFERENCE: AEWICA-X-2
;
; CURRENT APPLICATION NUMBER: US/10/029,386
;
; CURRENT FILING DATE: 2001-12-20
;
; NUMBER OF SEQ ID NOS: 34288
;
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;
; SEQ ID NO 31252
;
; LENGTH: 52
;

```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005254.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.61
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.65
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.58
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.74
; OTHER INFORMATION: SWISSPROT HIT: P55067, EVALUE 7.00e-28
US-10-029-386-31252
```

```
Query Match          49.4%; Score 84; DB 12; Length 52;
Best Local Similarity 42.3%; Pred. No. 0.0003;
Matches 11; Conservative 4; Mismatches 11; Indels 0; Gaps 0;
```

```
QY      2  CSSDWSSEYEGHCYKVFQSKTWTDAE 27
      | | | | | | | | | | | | | | | |
Db      4  CDRGWHKFGQHCYRYFAHRRAWEDAE 29
```

```
RESULT 15
US-10-241-220-82
; Sequence 82, Application US/10241220
; Publication No. US20030148406A1
; GENERAL INFORMATION:
; APPLICANT: Frantz, Gretchen
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Phillips, Heidi
; APPLICANT: Polakis, Paul
; APPLICANT: Spencer, Susan
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wu, Thomas
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF TUMOR
; FILE REFERENCE: P5010R1-US
; CURRENT APPLICATION NUMBER: US/10/241,220
; CURRENT FILING DATE: 2002-12-13
; NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 82
; LENGTH: 1321
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-241-220-82
```

```
Query Match          49.4%; Score 84; DB 12; Length 1321;
Best Local Similarity 42.3%; Pred. No. 0.0082;
Matches 11; Conservative 4; Mismatches 11; Indels 0; Gaps 0;
```

```
QY      2  CSSDWSSEYEGHCYKVFQSKTWTDAE 27
      | | | | | | | | | | | | | | | |
Db     1088 CDRGWHKFGQHCYRYFAHRRAWEDAE 1113
```

```
Search completed: December 8, 2003, 10:07:06
Job time : 11.9314 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 8, 2003, 12:41:24 ; Search time 1230.93 Seconds
(without alignments)
963.811 Million cell updates/sec

Title: US-09-938-114-3

Perfect score: 170
Sequence: 1 DCSDSWSSYEGHCYKVKFKQSKWTDAESF 29

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09938114/runat_08122003_091000_22849/app_query.fasta_1.725
-DB=GenEmbl -QWTF=fastap -SUPFIX=rge -MINMATCH=0 -1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPMT=PTO -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09938114 @CGN 1 1 3707 @runat_08122003_091000_22849 -NCPV=6 -ICPV=3
-NO.MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: gb_wa.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_on.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*

29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	163	95.9	733	5	AF176420 Deinagkis
2	155	91.2	497	5	AY091758 Deinagkis
3	155	91.2	656	5	AB036880 Deinagkis
4	148	87.1	688	5	AF540647 Deinagkis
5	136	80.0	458	5	AF176421 Deinagkis
6	134	78.8	483	6	AX427201 Sequence
7	134	78.8	664	5	AB019616 Agkistrod
8	133	78.2	478	5	AY091761 Deinagkis
9	133	78.2	592	5	AB036881 Deinagkis
10	133	78.2	698	5	D83332 Trimeresuru
11	132	77.6	369	5	AY293866 Deinagkis
12	132	77.6	454	6	AR259041 Sequence
13	132	77.6	478	5	AY091756 Deinagkis
14	132	77.6	574	5	AF350324 Deinagkis
15	127	74.7	634	5	AF197915 Gloydius
16	127	74.7	721	6	AX427204 Sequence
17	124	72.9	438	5	AF244901 Callosela
18	124	72.9	484	5	AY091760 Deinagkis
19	123	72.4	678	5	TFLETXA
20	123	72.4	683	5	AB046491 Trimeresu
21	121	71.2	375	5	AF387100 Deinagkis
22	119	70.0	453	6	AX427203 Sequence
23	119	70.0	583	5	AF125310 Gloydius
24	119	70.0	630	5	AF190827 Gloydius
25	116	68.2	478	5	AY091755 Deinagkis
26	116	68.2	685	5	AF102902 Deinagkis
27	116	68.2	738	5	Y16349 Crotalus du
28	115	67.6	456	6	AX427206 Sequence
29	115	67.6	544	6	AR259035 Sequence
30	114	67.1	487	5	AF463521 Deinagkis
31	114	67.1	565	5	AY268948 Echis car
32	114	67.1	619	5	AF540648 Deinagkis
33	111	65.3	690	6	AR026653 Sequence
34	111	65.3	690	6	AX404807 Sequence
35	111	65.3	690	6	BD143687 Protein h
36	110	64.7	580	6	AX427207 Sequence
37	103	60.6	408	5	AF244900 Callosela
38	103	60.6	432	6	AX427209 Sequence
39	103	60.6	508	5	AY091759 Deinagkis
40	103	60.6	666	5	AF540645 Deinagkis
41	102	60.0	441	5	AY099322 Protoboth
42	102	60.0	681	5	AY149340 Trimeresu
43	100	58.8	477	5	AY099321 Protoboth
44	99	58.2	632	5	AF125309 Gloydius
45	99	58.2	679	5	AB019615 Agkistrod

ALIGNMENTS

RESULT 1

AF176420 733 bp mRNA linear VRT 30-APR-2002
 LOCUS Deinagkistrodon acutus clone 2100490 agkisacutacin A chain mRNA, complete cds
 ACCESSION AF176420 AY091757
 VERSION AF176420.2 GI:20273043
 KEYWORDS
 SOURCE
 ORGANISM
 Deinagkistrodon acutus
 Deinagkistrodon acutus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea; Viperidae; Crotalinae; Deinagkistrodon.
 REFERENCE
 AUTHORS Yu.H.-X., Xiang,K.-J., and Liu,J.
 TITLE cDNA sequencing and analysis of eleven C-type lectin-like protein subunits from Agkistrodon acutus
 JOURNAL Sheng wu hua shueh yu sheng wu li hshueh pao (2002) In press
 REFERENCE
 AUTHORS Cheng,X., Qian,Y., Liu,Q., Li,B.X.Y., Ding,J., Xu,Z., Huang,W. and Liu,J.
 TITLE Agkisacutacin, a new fibrinolytic & anti-platelet protein from Agkistrodon acutus venom
 JOURNAL Unpublished
 REFERENCE
 AUTHORS Yu.H.-X., Xiang,K.-J., Wang,Y. and Liu,J.
 TITLE A chain of agkisacutacin from Deinagkistrodon acutus
 JOURNAL Unpublished
 REFERENCE
 AUTHORS Cheng,X., Liu,J., Li,B.X.Y. and Qian,Y.
 TITLE Direct Submission
 JOURNAL Submitted (08-AUG-1999) Dept. of Biochemistry and Molecular Biology, University of Science and Technology of China, School of Life-Science, Huangshan Road, Hefei, Anhui 230027, China
 REFERENCE
 AUTHORS Yu.H.-X., Xiang,K.-J., Wang,Y. and Liu,J.
 TITLE Direct Submission
 JOURNAL Submitted (23-APR-2002) Dept. of Biochemistry and Molecular Biology, University of Science and Technology of China, School of Life-Science, Huangshan Road, Hefei, Anhui 230027, China
 REMARK
 COMMENT On Apr 23, 2002 this sequence version replaced gi:6715112.
 FEATURES
 source
 1. /733
 /organism="Deinagkistrodon acutus"
 /mol_type="mRNA"
 /db_xref="taxon:36307"
 /clone="2100490"
 /tissue_type="venom"
 89. .547
 CDS
 /note="consists of two heterologous subunits (A and B chains) linked by an inter-subunit disulfide bond; c-type lectin family member"
 /codon_start=1
 /product="agkisacutacin A chain"
 /protein_id="AAF26286.2"
 /db_xref="GI:20273044"
 /translation="MGRIFFVSGLLVFLSLGTAADCSWSYEGHCHYKVPKQSK TWDAESFCTKQVNGHLVSISSGADFGVGLIAQKIKSAKHVIGLRAQNKQK SIEWSDGSSISYENWIEESKKGIVHETGFHWENFYCEQDPFVCEA"
 BASE COUNT 191 a 173 c 204 g 165 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2.57e-16 Length: 733
 Score: 163.00 Matches: 28
 Percent Similarity: 96.55% Conservative: 0
 Best Local Similarity: 96.55% Mismatches: 1
 Query Match: 95.88% Indels: 0
 DB: 5 Gaps: 0
 US-09-938-114-3 (1-29) x AF176420 (1-733)
 QY 1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
 DB 70 GATTGTTCTCTAGTTCCTCTCTATGAAGGCATTGCTACAGGCTTCAACAATCT 129
 QY 21 LysThrTrpThrAspAlaGluSerPhe 29
 DB 130 AAGACCTGGGACAGATGCAGAGAGCTTC 156
 RESULT 3
 AB036880
 LOCUS
 DEFINITION
 AB036880
 VERSION
 AB036880.1 GI:8980618

Db 158 GATTGTTCTCTAGTTCCTCTCTATGAAGGCATTGCTACAGGCTTCAACAATCT 217
 QY 21 LysThrTrpThrAspAlaGluSerPhe 29
 Db 218 AAGACCTGGGACAGATGCAGAGAGCTTC 244
 RESULT 2
 AY091758
 LOCUS Deinagkistrodon acutus clone 2100590 ACF 1/2 A-chain mRNA, complete cds
 DEFINITION
 ACCESSION AY091758.1 GI:20562936
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Deinagkistrodon acutus
 Deinagkistrodon acutus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea; Viperidae; Crotalinae; Deinagkistrodon.
 REFERENCE
 AUTHORS Yu.H., Xiang,K., Wang,Y. and Liu,J.
 TITLE A chain of ACF 1/2 from Deinagkistrodon acutus
 JOURNAL Unpublished
 REFERENCE
 AUTHORS Yu.H., Xiang,K., Wang,Y. and Liu,J.
 TITLE Direct Submission
 JOURNAL Submitted (22-MAR-2002) Dept. of Biochemistry and Molecular Biology, University of Science and Technology of China, School of Life Science, Huangshan Road, Hefei, Anhui 230027, P.R. China
 FEATURES
 source
 1. /497
 /organism="Deinagkistrodon acutus"
 /mol_type="mRNA"
 /db_xref="taxon:36307"
 /clone="2100590"
 1. 459
 /note="C-type lectin family member"
 /codon_start=1
 /product="ACF 1/2 A-chain"
 /protein_id="AAM22786.1"
 /db_xref="GI:20562937"
 /translation="MGRIFFVSGLLVFLSLGTAADCSWSYEGHCHYKVPKQSK TWDAESFCTKQVNGHLVSISSGADFGVGLIAQKIKSAKHVIGLRAQNKQK SIEWSDGSSISYENWIEESKKGIVHETGFHWENFYCEQDPFVCEA"
 BASE COUNT 135 a 102 c 143 g 117 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3.12e-15 Length: 497
 Score: 155.00 Matches: 26
 Percent Similarity: 89.66% Conservative: 0
 Best Local Similarity: 89.66% Mismatches: 3
 Query Match: 91.18% Indels: 0
 DB: 5 Gaps: 0
 US-09-938-114-3 (1-29) x AY091758 (1-497)
 QY 1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
 Db 70 GATTGTTCTCTAGTTCCTCTCTATGAAGGCATTGCTACAGGCTTCAACAATCT 129
 QY 21 LysThrTrpThrAspAlaGluSerPhe 29
 Db 130 AAGACCTGGGACAGATGCAGAGAGCTTC 156
 RESULT 3
 AB036880
 LOCUS
 DEFINITION
 AB036880
 VERSION
 AB036880.1 GI:8980618


```

KEYWORDS    anticoagulant protein A.
SOURCE      Deinagkistrodon acutus
ORGANISM    Deinagkistrodon acutus
REFERENCE   1
AUTHORS     Tani, A., Ogawa, T., Nose, T., Nikandrov, N.N., Deshimaru, M.,
            Chijiwa, T., Chang, C.C., Fukumaki, Y. and Ohno, M.
TITLE       Characterization, primary structure and molecular evolution of
            anticoagulant protein from Agkistrodon actus venom
JOURNAL     Toxicon 40 (6), 803-813 (2002)
MEDLINE     21975857
REFERENCE   2 (bases 1 to 656)
AUTHORS     Ogawa, T. and Tani, A.
TITLE       Direct Submission
JOURNAL     Submitted (06-JAN-2000) Tomohisa Ogawa, Tohoku University, Graduate
            School of Agricultural Science; 1-1, Tsutsumidori Amamiyamachi,
            Aoba-ku, Sendai 981-8555, Japan [E-mail:ogawa@biochem.tohoku.ac.jp,
            URL:http://www.agri.tohoku.ac.jp/hooso/, Tel:81-22-717-8808,
            Fax:81-22-717-8807]
FEATURES    Location/Qualifiers
             source
               1..656
               /organism="Deinagkistrodon acutus"
               /mol_type="mRNA"
               /db_xref="taxon:36307"
               /tissue_type="venom gland"
             gene
               1..656
               /gene="acp-a"
               /codon_start=1
               /product="anticoagulant protein A"
               /protein_id="BAA99281.1"
               /db_xref="GI:8980619"
               /translation="MGRFIFVSGLLVVLVLSGTAADCPSSWSSEYGHCVKAPKQSK
               TWADAESFCTQVNGHSLVSESGEADFVAHLIAQIKSAKIHVIGLKNQKQK
               SIENWDSGSSISYENWIEESKCLGVHKAQFRKWFNYCFQRDPFVCEA"
             sig_peptide
               50..118
               /gene="acp-a"
BASE COUNT  173 a 148 c 185 g 150 t
ORIGIN
Alignment Scores:
Pred. No.:    4.25e-15      Length:    656
Score:        155.00      Matches:    26
Percent Similarity:  89.66%      Conservative:  0
Best Local Similarity:  89.66%      Mismatches:  3
Query Match:   91.18%      Indels:    0
DB:           5           Gaps:    0

US-09-938-114-3 (1-29) x AB036880 (1-656)
Qy  1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
Db  119 GATTGTTCTCTAGTGGTCTCTCTATGAGGGCATTTGCTATAGGCGCTTCAAAACAATCT 178
Qy  21 LysThrTrpThrAspAlaGluSerPhe 29
Db  179 AAGACCTGGGCAGATGCAGAGAGCTTC 205

RESULT 4
AF540647
LOCUS      AF540647
DEFINITION Deinagkistrodon acutus agglucetin-beta 1 subunit precursor, mRNA,
            complete cds.
ACCESSION AF540647
VERSION   AF540647.1 GI:23321262
KEYWORDS
SOURCE     Deinagkistrodon acutus
ORGANISM   Deinagkistrodon acutus
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;

Viperidae; Crotalinae; Deinagkistrodon.
1 (bases 1 to 688)
Wang, W.-J., Ling, Q.-D. and Huang, T.-F.
A novel tetrameric venom protein, agglucetin from Agkistrodon
acutus, acts as a glycoprotein Ib agonist
Thromb. Haemost. 86 (4), 1077-1086 (2001)
21542539
JOURNAL MEDLINE
PUBMED 11686327
REFERENCE 2 (bases 1 to 688)
AUTHORS Wang, W.-J., Ling, Q.-D. and Huang, T.-F.
TITLE Molecular structure and functional characterization of agglucetin,
a tetrameric glycoprotein Ib-binding protein, from Formosan pit
viper
Unpublished
JOURNAL REFERENCE 3 (bases 1 to 688)
AUTHORS Wang, W.-J., Ling, Q.-D. and Huang, T.-F.
TITLE Direct Submission
JOURNAL Submitted (23-AGU-2002) Department of Nursing, Chang Gung Institute
of Technology, 261, Wen-Hua 1 Rd., Kwei-Shan, Tao-Yuan, Taiwan
333-03, R.O.C.
FEATURES Location/Qualifiers
             source
               1..688
               /organism="Deinagkistrodon acutus"
               /mol_type="mRNA"
               /db_xref="taxon:36307"
               /note="platelet agglutination inducer; glycoprotein Ib
               agonist"
               /codon_start=1
               /product="agglucetin-beta 1 subunit precursor"
               /protein_id="AA23126.1"
               /db_xref="GI:23321263"
               /translation="MGRVIFVSGLLVVLVLSGTAADCPSEWSSEYGHCVKAPKQSK
               TWADAERFCTQKHGSHLAFHSSEADFVVTLPSTLKLTLVWGLKNINWGCYKWK
               SDGPKLDYKDWRQEFCLVSRVYNNWLSMDGTTYSFVCKFQA"
             sig_peptide
               61..129
             mat_peptide
               130..498
BASE COUNT  181 a 173 c 172 g 162 t
ORIGIN
Alignment Scores:
Pred. No.:    5.8e-14      Length:    688
Score:        148.00      Matches:    24
Percent Similarity:  86.21%      Conservative:  1
Best Local Similarity:  82.76%      Mismatches:  4
Query Match:   87.06%      Indels:    0
DB:           5           Gaps:    0

US-09-938-114-3 (1-29) x AF540647 (1-688)
Qy  1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
Db  130 GATTGTTCTCTAGTGGTCTCTCTATGAGGGCATTTGCTACAGGCGCTTCAAAACAATCT 189
Qy  21 LysThrTrpThrAspAlaGluSerPhe 29
Db  190 AAGACCTGGGCAGATGCAGAGAAATTC 216

RESULT 5
AF176421
LOCUS      AF176421
DEFINITION Deinagkistrodon acutus agkisacutacin B chain mRNA, complete cds.
ACCESSION AF176421
VERSION   AF176421.1 GI:6715114
KEYWORDS
SOURCE     Deinagkistrodon acutus
ORGANISM   Deinagkistrodon acutus
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Deinagkistrodon.
            1 (bases 1 to 458)
            Cheng, X., Qian, Y., Liu, Q., Li, B. X. Y., Ding, J., Xu, Z., Huang, W. and
            Huang, W. and

```

```

Liu, J.
Agkisacutacin, a new fibrinolytic & anti-platelet protein from
Agkistrodon acutus venom
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 458)
AUTHORS Cheng, X., Liu, J., Li, B. X. Y. and Qian, Y.
TITLE Direct Submission
JOURNAL Submitted (08-AUG-1999) Dept. of Biochemistry and Molecular
Biology, University of Science and Technology of China, School of
Life-Science, Huangshan Road, Hefei, Anhui 230027, China
FEATURES
source
Location/Qualifiers
1..458
/organism="Deinagkistrodon acutus"
/mol_type="mRNA"
/db_xref="taxon:36307"
/tissue_type="venom"
3..443
/note="agkisacutacin consists of two heterologous subunits
(A and B chains) linked by an inter-subunit disulfide
bond"
/codon_start=1
/product="agkisacutacin B chain"
/protein_id="AAF26287.1"
/db_xref="GI:6715115"
/translation="MGRIFFVSGLLVFLSLGTAAADCPDSWSSYEGHCYKPFDEPK
TWDAKEFCTQKHGSHLASHFSSEADPVLTPTSLKTLVLWGLKNIWGCYWKW
SDGKLDYKDWREDFECLVSRIVNEMLSMDCGTCISFVCRFQA"
BASE COUNT 118 a 109 c 116 g 115 t
ORIGIN
1
2 98e-12 Length: 458
Score: 136.00 Matches: 22
Percent Similarity: 79.31% Conservativeness: 1
Best Local Similarity: 75.86% Mismatches: 6
Query Match: 80.00% Indels: 0
DB: 5 Gaps: 0
US-09-938-114-3 (1-29) x AF176421 (1-458)
Qy 1 AspCysSerSerAspTyrTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
|||||
72 GATTGTCCCTCTGATTGGTCTCTCTATGAAGGGCATTGCTCAAGCCCTTCGATGAACCT 131
|||||
Qy 21 LysThrTrpThrAspAlaGluSerPhe 29
|||||
132 AAGACCTGGCAGATGCAGAGAAATTC 158
|||||
RESULT 6
LOCUS AX427201 483 bp DNA linear PAT 18-JUN-2002
DEFINITION Sequence 1 from Patent WO0214364.
ACCESSION AX427201
VERSION AX427201.1 GI:21530563
KEYWORDS
SOURCE
ORGANISM
Sistrurus miliarius (pigmy rattlesnake)
Sistrurus miliarius
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
Viperidae; Crotalinae; Sistrurus.
1
Sheppard, P. O. and Bishop, P. D.
Rattlesnake venom gland proteins
Patent: WO 0214364-A 1 21-FEB-2002;
ZymoGenetics, Inc. (US)
Location/Qualifiers
1..483
/organism="Sistrurus miliarius"
/mol_type="genomic DNA"
/db_xref="taxon:8758"
<3..458
/note="Zsnk2"
/codon_start=1
CDS
Alignment Scores:
Pred. No.: 9.35e-12 Length: 664
Score: 134.00 Matches: 483
Percent Similarity: 82.76% Conservativeness: 2
Best Local Similarity: 75.86% Mismatches: 5
Query Match: 78.82% Indels: 0
DB: 6 Gaps: 0
US-09-938-114-3 (1-29) x AX427201 (1-483)
Qy 1 AspCysSerSerAspTyrTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
|||||
60 GATTGTCCCTCTGATTGGTCTCTCTATGATCAGCATGCTACAGGTCTTCAGTGAATC 119
|||||
Qy 21 LysThrTrpThrAspAlaGluSerPhe 29
|||||
120 AAACCTGGGATGATGCAGAGAGTTTC 146
|||||
RESULT 7
LOCUS AB019616 664 bp mRNA linear VRT 14-NOV-1998
DEFINITION Agkistrodon blomhoffi mRNA for mamushigin beta, complete cds.
ACCESSION AB019616
VERSION AB019616.1 GI:3882118
KEYWORDS mamushigin beta.
SOURCE Agkistrodon blomhoffi
ORGANISM Agkistrodon blomhoffi
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
Viperidae; Crotalinae; Agkistrodon.
1 (sites)
Sakurai, Y., Fujimura, Y., Kokubo, T., Imamura, K., Kawasaki, T.,
Handa, M., Suzuki, M., Matsui, T., Titani, K. and Yoshioka, A.
The cDNA cloning and molecular characterization of a snake venom
platelet glycoprotein Ib-binding protein, mamushigin, from
Agkistrodon halsys blomhoffii venom
Thromb. Haemost. 79 (6), 1199-1207 (1998)
98319530
PUBMED 9657448
REFERENCE 2 (bases 1 to 664)
AUTHORS Sakurai, Y. and Fujimura, Y.
TITLE Direct Submission
JOURNAL Submitted (10-NOV-1998) Yoshihiko Sakurai, Nara Medical University,
Department of Blood Transfusion Medicine; Shijo-cho 840, Kashihara,
Nara 634-8522, Japan (E-mail:ysakurai@nmu-gw.cc.naramed-u.ac.jp,
Tel:81-744-22-3051(ex.3288), Fax:81-744-29-0771)
FEATURES
source
Location/Qualifiers
1..664
/organism="Agkistrodon blomhoffi"
/mol_type="mRNA"
/db_xref="taxon:61300"
/note="synonym:Agkistrodon halsys blomhoffii"
71..511
/codon_start=1
/product="mamushigin beta"
/protein_id="BA034425.1"
/db_xref="GI:3882119"
/translation="MGRFIFLSGLLVFLVFLVSLGTADCPDSWSSYEGHCYKRVFOKEM
TWDAKEFCTQKHSHLASHFSSEADPVVMTPIKLYDFVWGLNINWECVW
TDGRLSHNMTITSECTIAAKTQNLRSRCSRTYNVVCCKFE"
BASE COUNT 174 a 153 c 173 g 164 t
ORIGIN
Alignment Scores:
Pred. No.: 9.35e-12 Length: 664

```

```

/protein_id="CAD35498.1"
/db_xref="GI:21530564"
/translation="IFVSGLLVFLVFLSLGTADCPDSWSSYDOHCYKVFSELKTTWDD
AESFCYQHRDSRLASHSSEEAFLVGLKSLQTLKFTSMWIGLKLWKECKWQSSDDT
KLDYKAWTRRPYCTVMVVKTDRIEWFNRGCEKTVSFVCKFOARSGDPAV"
BASE COUNT 129 a 112 c 123 g 119 t
ORIGIN
Alignment Scores:
Pred. No.: 6.57e-12 Length: 483
Score: 134.00 Matches: 22
Percent Similarity: 82.76% Conservativeness: 2
Best Local Similarity: 75.86% Mismatches: 5
Query Match: 78.82% Indels: 0
DB: 6 Gaps: 0
US-09-938-114-3 (1-29) x AX427201 (1-483)
Qy 1 AspCysSerSerAspTyrTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
|||||
60 GATTGTCCCTCTGATTGGTCTCTCTATGATCAGCATGCTACAGGTCTTCAGTGAATC 119
|||||
Qy 21 LysThrTrpThrAspAlaGluSerPhe 29
|||||
120 AAACCTGGGATGATGCAGAGAGTTTC 146
|||||
RESULT 7
LOCUS AB019616 664 bp mRNA linear VRT 14-NOV-1998
DEFINITION Agkistrodon blomhoffi mRNA for mamushigin beta, complete cds.
ACCESSION AB019616
VERSION AB019616.1 GI:3882118
KEYWORDS mamushigin beta.
SOURCE Agkistrodon blomhoffi
ORGANISM Agkistrodon blomhoffi
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
Viperidae; Crotalinae; Agkistrodon.
1 (sites)
Sakurai, Y., Fujimura, Y., Kokubo, T., Imamura, K., Kawasaki, T.,
Handa, M., Suzuki, M., Matsui, T., Titani, K. and Yoshioka, A.
The cDNA cloning and molecular characterization of a snake venom
platelet glycoprotein Ib-binding protein, mamushigin, from
Agkistrodon halsys blomhoffii venom
Thromb. Haemost. 79 (6), 1199-1207 (1998)
98319530
PUBMED 9657448
REFERENCE 2 (bases 1 to 664)
AUTHORS Sakurai, Y. and Fujimura, Y.
TITLE Direct Submission
JOURNAL Submitted (10-NOV-1998) Yoshihiko Sakurai, Nara Medical University,
Department of Blood Transfusion Medicine; Shijo-cho 840, Kashihara,
Nara 634-8522, Japan (E-mail:ysakurai@nmu-gw.cc.naramed-u.ac.jp,
Tel:81-744-22-3051(ex.3288), Fax:81-744-29-0771)
FEATURES
source
Location/Qualifiers
1..664
/organism="Agkistrodon blomhoffi"
/mol_type="mRNA"
/db_xref="taxon:61300"
/note="synonym:Agkistrodon halsys blomhoffii"
71..511
/codon_start=1
/product="mamushigin beta"
/protein_id="BA034425.1"
/db_xref="GI:3882119"
/translation="MGRFIFLSGLLVFLVFLVSLGTADCPDSWSSYEGHCYKRVFOKEM
TWDAKEFCTQKHSHLASHFSSEADPVVMTPIKLYDFVWGLNINWECVW
TDGRLSHNMTITSECTIAAKTQNLRSRCSRTYNVVCCKFE"
BASE COUNT 174 a 153 c 173 g 164 t
ORIGIN
Alignment Scores:
Pred. No.: 9.35e-12 Length: 664

```

Score: 134.00 Matches: 21
Percent Similarity: 82.76% Conservative: 3
Best Local Similarity: 72.41% Mismatches: 5
Query Match: 78.82% Indels: 0
DB: 5 Gaps: 0

US-09-938-114-3 (1-29) x AB019616 (1-664)

QY 1 AspCysSerSerAspTrpSerSerTyGluGlyHisCysTyrlsValPheLysGlnSer 20
140 GATTGTCCTCTGATTGGCTCTCTATCAAGGGCATTGCTACAGGCTTCCAAAGAG 199
21 LysThrTrpThrAspAlaGluSerPhe 29
200 ATGACCTGGGAAGATGCAGAAATTC 226

RESULT 8
AY091761
LOCUS
DEFINITION
Deinagkistrodon acutus clone 2101 ACF 1/2 B-chain mRNA, complete cds
ACCESSION
AY091761.1 GI:20562942
VERSION
AY091761.1
KEYWORDS
Deinagkistrodon acutus
SOURCE
Deinagkistrodon acutus
ORGANISM
Deinagkistrodon acutus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea; Viperidae; Crotalinae; Deinagkistrodon.
REFERENCE
YU, H., XIANG, K., WANG, Y. and LIU, J.
AUTHORS
YU, H., XIANG, K., WANG, Y. and LIU, J.
TITLE
B chain of ACF 1/2 from Deinagkistrodon acutus
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 478)
AUTHORS
YU, H., XIANG, K., WANG, Y. and LIU, J.
TITLE
Direct Submission
JOURNAL
Submitted (22-MAR-2002) Dept. of Biochemistry and Molecular Biology, University of Science and Technology of China, School of Life Science, Huangshan Road, Hefei, Anhui 230027, P.R. China

FEATURES
source
1. .478
/organism="Deinagkistrodon acutus"
/mol_type="mRNA"
/db_xref="taxon:36307"
/clone="2101"
1. .441
/notes="C-type lectin family member"
/codon_start=1
/product="ACF 1/2 B-chain"
/protein_id="AAM22789.1"
/db_xref="GI:20562943"
/translation="MGRFIFVSGLLVFLSLSGTGADCPDSSVEGHCKYKPFNEPK
SNAAENFCTQOHTGSHLVSPSTEEADPVKLAFTDYGIFFWGLSKINQCNQW
SNAAMLYTDWAEESYCVYFKSTNNKWSITCRMIANFVCEFOA"
BASE COUNT 127 a 107 c 125 g 119 t
ORIGIN

Alignment Scores:
Pred. No.: 9.36e-12 Length: 478
Score: 133.00 Matches: 21
Percent Similarity: 79.31% Conservative: 2
Best Local Similarity: 72.41% Mismatches: 6
Query Match: 78.24% Indels: 0
DB: 5 Gaps: 0

US-09-938-114-3 (1-29) x AY091761 (1-478)

QY 1 AspCysSerSerAspTrpSerSerTyGluGlyHisCysTyrlsValPheLysGlnSer 20
70 GATTGTCCTCTGATTGGCTCTCTATCAAGGGCATTGCTACAGGCTTCCAAAGAG 129
21 LysThrTrpThrAspAlaGluSerPhe 29
200 ATGACCTGGGAAGATGCAGAAATTC 226

RESULT 10
TFLFIXB
LOCUS

Db 130 AAAAATGGCGGATGCAGAAATTC 156

RESULT 9
AB036881
LOCUS
DEFINITION
Deinagkistrodon acutus acp-b mRNA for anticoagulant protein-B, complete cds.
ACCESSION
AB036881.1 GI:12060180
VERSION
AB036881.1
KEYWORDS
anticoagulant protein-B.
SOURCE
Deinagkistrodon acutus
ORGANISM
Deinagkistrodon acutus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea; Viperidae; Crotalinae; Deinagkistrodon.
REFERENCE
1
AUTHORS
Tani, A., Ogawa, T., Nose, T., Nikandrov, N.N., Deshimaru, M., Chijiwa, T., Chang, C.C., Fukumaki, Y. and Ohno, M.
TITLE
Characterization, primary structure and molecular evolution of anticoagulant protein from Agkistrodon actus venom
JOURNAL
Toxicon 40 (6), 803-813 (2002)
MEDLINE
21975857
REFERENCE
2 (bases 1 to 592)
AUTHORS
Ogawa, T. and Tani, A.
TITLE
Direct Submission
JOURNAL
Submitted (14-JAN-2000) Tomohisa Ogawa, Tohoku University, Graduate School of Agricultural Science, 1-1, Tsutsumidori Anamiyamachi, Aoba-ku, Sendai 981-8555, Japan (E-mail:ogawa@biochem.tohoku.ac.jp, URL:http://www.agri.tohoku.ac.jp/hoozo/, Tel:81-22-717-8808, Fax:81-22-717-8807)

FEATURES
source
1. .592
/organism="Deinagkistrodon acutus"
/mol_type="mRNA"
/db_xref="taxon:36307"
/tissue_type="venom gland"
1. .592
/genes="acp-b"
5. .445
/genes="acp-b"
/codon_start=1
/product="anticoagulant protein-B"
/protein_id="BAB20441.1"
/db_xref="GI:12060181"
/translation="MGRFIFVSGLLVFLSLSGTGADCPDSSVEGHCKYKPFNEPK
NWADAEFCTQOHTGSHLVSPSTEEADPVKLAFTDYGIFFWGLSKINQCNQW
SNAAMLYTDWAEESYCVYFKSTNNKWSITCRMIANFVCEFOA"
sig_peptide 5. .73
BASE COUNT 149 a 146 c 149 g 148 t
ORIGIN

Alignment Scores:
Pred. No.: 1.19e-11 Length: 592
Score: 133.00 Matches: 21
Percent Similarity: 79.31% Conservative: 2
Best Local Similarity: 72.41% Mismatches: 6
Query Match: 78.24% Indels: 0
DB: 5 Gaps: 0

US-09-938-114-3 (1-29) x AB036881 (1-592)

QY 1 AspCysSerSerAspTrpSerSerTyGluGlyHisCysTyrlsValPheLysGlnSer 20
74 GATTGTCCTCTGATTGGCTCTCTATCAAGGGCATTGCTACAGGCTTCCAAAGAG 133

Db 21 LysThrTrpThrAspAlaGluSerPhe 29
134 AAAAATGGCGGATGCAGAAATTC 160

RESULT 10
TFLFIXB
LOCUS

TFLFIXB 698 bp mRNA linear VRT 06-FEB-1999

DEFINITION Trimeresurus flavoviridis mRNA for factor IX/factor X binding protein B chain, complete cds.
ACCESSION D83332
VERSION D83332.1 GI:1402641
KEYWORDS factor IX/factor X binding protein B chain.
SOURCE Trimeresurus flavoviridis
ORGANISM Trimeresurus flavoviridis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Lepidodonta; Squamata; Scleroglossa; Serpentes; Colubroidea; Viperidae; Crotalinae; Trimeresurus.
REFERENCE 1 (bases 1 to 698)
AUTHORS Matsuzaki, R., Yoshiara, E., Yamada, M., Shima, K., Atoda, H. and Morita, T.
TITLE cDNA cloning of IX/X-BP, a heterogeneous two-chain anticoagulant protein from snake venom
JOURNAL Biochem. Biophys. Res. Commun. 220 (2), 382-387 (1996)
MEDLINE 96184662
PUBMED 8645314
REFERENCE 2 (bases 1 to 698)
AUTHORS Morita, T.
TITLE Direct Submission
JOURNAL Biochemistry; 1-22-1 Yatocho, Tanashi, Tokyo 188, Japan
 Submitted (02-FEB-1996) Takashi Morita, Meiji College of Pharmacy,
 Biochemistry; 1-22-1 Yatocho, Tanashi, Tokyo 188, Japan
 (Tel:0424-21-0101(ex.429), Fax:0424-21-1489)
FEATURES Location/Qualifiers
 source
 1..698
 /organism="Trimeresurus flavoviridis"
 /mol_type="mRNA"
 /db_xref="taxon:88087"
 1..73
 74..514
 /codon_start=1
 /product="factor IX/factor X binding protein B chain"
 /protein_id="BA011888.1"
 /db_xref="GI:1402642"
 /translation="MGRFIFMGLVFLSLGTAADCPDMSVYEGHCYKPFSEPK
 NWADAENFCTQOAGHGLVFSQSEADFVVKLAFQTFGHSIFPMGLSNVWNCNWQW
 SNAMLEYKAAEESYCVYFKSTNNKRSACRMQAQVCEFOA"
 515..698
 polyA_signal 679..684
3'UTR 167 a 185 c 177 g 169 t
BASE COUNT
ORIGIN
 Alignment Scores:
 Pred. No.: 1.42e-11 Length: 698
 Score: 133.00 Matches: 21
 Percent Similarity: 79.31% Conservative: 2
 Best Local Similarity: 72.41% Mismatches: 6
 Query Match: 78.24% Indels: 0
 DB: 5 Gaps: 0
 US-09-938-114-3 (1-29) x TFLFIXB (1-698)
Qy 1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 143 GATTGCTCCTGATTGCTCTTATGAGGCGATTGCTACAGCCCTTCAGTGAACCC 202
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 21 LysThrTrpThrAspAlaGluSerPhe 29
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 203 AAAAACTGGCGCGATGACAGAAATTC 229
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 11
LOCUS AY293866 369 bp DNA linear VRT 09-JUN-2003
DEFINITION Deinagkistrodon acutus agkisasin-b gene, partial cds.
ACCESSION AY293866
VERSION AY293866.1 GI:31559054
KEYWORDS
SOURCE Deinagkistrodon acutus
ORGANISM Deinagkistrodon acutus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Lepidodonta; Squamata; Scleroglossa; Serpentes; Colubroidea;
 Viperidae; Crotalinae; Deinagkistrodon.

REFERENCE 1 (bases 1 to 369)
AUTHORS Zha, X.D., Ren, B., Liu, J. and Xu, K.S.
TITLE Genomic DNA sequence of b chain of Agkisasin, a C-type lectin-like protein from Agkistrodon acutus venom, and its evolutionary significance
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 369)
AUTHORS Zha, X.D., Ren, B., Liu, J. and Xu, K.S.
TITLE Direct Submission
JOURNAL Submitted (07-MAY-2003) Biology, School of Life Science, Anhui University, 3 Feixi Road, Hefei 230039, China
FEATURES Location/Qualifiers
 source
 1..369
 /organism="Deinagkistrodon acutus"
 /mol_type="genomic DNA"
 /db_xref="taxon:36307"
 <1..369
 /product="agkisasin-b"
 <1..369
 /notes="C-type lectin-like protein"
 /codon_start=1
 /product="agkisasin-b"
 /protein_id="AAP50528.1"
 /db_xref="GI:31559055"
 /translation="DCPSDMSVYEGHCYKPFNKLKNWADAENFCTQOHTGSHLVSFOS
 TEADSVVKLAFQTFYGFPMGLSNVWNCNWQSWNAAMKYTDMAEESYCVYFKST
 NNKRSITCRMIANFVCEFOA"
 107 a 83 c 88 g 91 t
BASE COUNT
ORIGIN
 Alignment Scores:
 Pred. No.: 1.01e-11 Length: 369
 Score: 132.00 Matches: 21
 Percent Similarity: 79.31% Conservative: 2
 Best Local Similarity: 72.41% Mismatches: 6
 Query Match: 77.65% Indels: 0
 DB: 5 Gaps: 0
 US-09-938-114-3 (1-29) x AY293866 (1-369)
Qy 1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 GATTGCTCCTGATTGCTCTTATGAGGCGATTGCTACAGCCCTTCATGAATG 60
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 21 LysThrTrpThrAspAlaGluSerPhe 29
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 AAAAACTGGCGCGATGACAGAAATTC 87
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 12
LOCUS AR259041 454 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 12 from patent US 6489451.
ACCESSION AR259041
VERSION AR259041.1 GI:27309526
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 454)
AUTHORS Li, B.X. and Cheng, X.
TITLE Antithrombotic enzyme from the snake venom of agkistrodon acutus
JOURNAL Patent: US 6489451-A 12 03-DEC-2002;
FEATURES Location/Qualifiers
 source
 1..454
 /organism="unknown"
BASE COUNT 114 a 107 c 114 g 109 t 10 others
ORIGIN
 Alignment Scores:
 Pred. No.: 1.27e-11 Length: 454
 Score: 132.00 Matches: 21
 Percent Similarity: 79.31% Conservative: 2
 Best Local Similarity: 72.41% Mismatches: 6

Query Match:	77.65%	Indels:	0
DB:	6	Gaps:	0
US-09-938-114-3 (1-29) x AR259041 (1-454)			
QY	1 AspCysSerSerAspTyrSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20		
DB	72 GATTGTCCTCTGATGGTCTCTCTATGAAGGCATTGCTACAAGCCCTTCGATGAACCT 131		
QY	21 LysThrTyrThrAspAlaGluSerPhe 29		
DB	132 AAGACCTGGCAGATGCAGAGAAATTC 158		
RESULT 13			
AY091756	478 bp mRNA linear VRT 13-MAY-2002		
LOCUS	Deinagkistrodon acutus clone 2100488 agkisacutacin B-chain mRNA, complete cds.		
DEFINITION	AY091756		
ACCESSION	AY091756.1 GI:20562934		
VERSION			
KEYWORDS			
SOURCE	Deinagkistrodon acutus		
ORGANISM	Deinagkistrodon acutus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea; Viperidae; Crotalinae; Deinagkistrodon.		
AUTHORS	Yu, H., Xiang, K., Wang, Y. and Liu, J.		
TITLE	B chain of agkisacutacin from Deinagkistrodon acutus		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 478)		
AUTHORS	Yu, H., Xiang, K., Wang, Y. and Liu, J.		
TITLE	Direct Submission		
JOURNAL	Submitted (22-MAR-2002) Dept. of Biochemistry and Molecular Biology, University of Science and Technology of China, School of Life Science, Huangshan Road, Hefei, Anhui 230027, P.R. China		
FEATURES	Location/Qualifiers		
source	1..478		
	/organism="Deinagkistrodon acutus"		
	/mol_type="mRNA"		
	/db_xref="taxon:36307"		
	/tissue_type="venom gland"		
	/lab_host="Escherichia coli JM109"		
	1..375		
	/note="IX/X-Bp family member; coagulant and anticoagulant"		
	/codon_start=1		
	/product="agkisacin-b"		
	/protein_id="AAK26430.1"		
	/db_xref="GI:13445904"		
	/translation="MDCPSDMSYEGHCYKPFENELKNWADAENFCTQOHTGSHLYSFG STEEDFVVKLAFOTFDYGI PMWGLSNIQWQNSNAMLKYTDWAEESYCVTFKS TNKRSITCRMIANFVCEQA"		
BASE COUNT	170 a	134 c	133 g 137 t
ORIGIN			
Alignment Scores:			
Pred. No.:	1..65e-11	Length:	574
Score:	132.00	Matches:	21
Percent Similarity:	79.31%	Conservative:	2
Best Local Similarity:	72.41%	Mismatches:	6
Query Match:	77.65%	Indels:	0
DB:	5	Gaps:	0
US-09-938-114-3 (1-29) x AF350324 (1-574)			
QY	1 AspCysSerSerAspTyrSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20		
DB	4 GATTGTCCTCTGATGGTCTCTCTATGAAGGCATTGCTACAAGCCCTTCGATGAACCTG 63		
QY	21 LysThrTyrThrAspAlaGluSerPhe 29		
DB	64 AAAAAGCTGGCGGATGCAGAGAAATTC 90		
RESULT 15			
AF197915	634 bp mRNA linear VRT 01-NOV-2000		
LOCUS	Gloydius halyx halyxin B-chain precursor, mRNA, complete cds.		
DEFINITION	AF197915		
ACCESSION	AF197915.1 GI:11066255		
VERSION			
KEYWORDS			
SOURCE	Gloydius halyx (halyx viper)		
ORGANISM	Gloydius halyx		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea; Viperidae; Crotalinae; Gloydius.		
AUTHORS	Koo, B.H., Sohn, Y.D., Kim, D.S., Jang, Y.S. and Chung, K.H.		
TITLE	A Novel Coagulation Factor Xa Inhibitor from Korean Snake (Agkistrodon halyx) venom		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 634)		
AUTHORS	Sohn, Y.D., Koo, B.H., Kim, D.S., Jang, Y.S. and Chung, K.H.		

TITLE Direct Submission
JOURNAL Submitted (22-OCT-1999) Cardiovascular center, Yonsei University
College of Med., 134 Shinchon-dong, Seoul 120-752, Republic of
Korea

FEATURES

source

Location/Qualifiers

1. .634
/organism="Gloydius halys"
/mol_type="mRNA"
/db_xref="taxon:8714"
/tissue_type="venom"

24. .464

/note="coagulation factor Xa inhibitor"

/codon_start=1

/product="halyxin B-chain precursor"

/protein_id="AAG28522.1"

/db_xref="GI:11066256"

/translation="MGRFIFLSFGLVVLVLSLGTGADCPGWSVYEGHCYKPFNEQK
NWADAENFCTQHTGHLVSPHTEADFVVKLAFONFGHGIWMGLSNVWNCQSQW
SSAAKLKYEAWAESYCVYFKSTNNKRSRACRMEAYFVCEFQA"

24. .92

sig_peptide

93. .461

mat_peptide

/product="halyxin B-chain"

BASE COUNT

180 a 142 c 157 g 155 t

ORIGIN

Alignment Scores: 1.15e-10 Length: 634

Pred. No.: 127.00 Matches: 20

Score: 127.00 Matches: 20

Percent Similarity: 75.86% Conservative: 2

Best Local Similarity: 68.97% Mismatches: 7

Query Match: 74.71% Indels: 0

DB: 5 Gaps: 0

US-09-938-114-3 (1-29) x AF197915 (1-634)

Qy 1 AspCysSerSerAspTrpSerSertyrGluGlyHisCysTyrLysValPheLysGlnSer 20

Db 93 GATTGTCCTCTGGTGGTGGTCTCTCTATGAGGGCATTCCTACCAACCATTTAATGACAA 152

Qy 21 LysThrTrpThrAspAlaGluSerPhe 29

Db 153 AAAAATTGGCGGATGCAGAGATTTC 179

Search completed: December 8, 2003, 16:08:45

Job time : 1231.93 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 8, 2003, 12:39:59 ; Search time 99.5943 Seconds
(without alignments)
786.026 Million cell updates/sec

Title: US-09-938-114-3

Perfect score: 170

Sequence: 1 DCSWSYEGHCYKVPKQKWTDAESF 29

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DBV=xlh
-Q=/cgn2_1/USPTO.spool/US0938114/runat_08122003_091000_22839/app_query.fasta_1.725
-DB=N_geneseq 19Jun03 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS-human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0938114 @CGN 1.1.0 @runat_08122003_091000_22839 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 19Jun03.+

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	134	78.8	483	24	AAD32053	Pigmy rattlesnake
2	132	77.6	454	25	ABX93674	CDNA encoding Dein
3	127	74.7	633	22	AAI71877	Snake venom blood
4	127	74.7	721	24	AAD32055	Pigmy rattlesnake
5	119	70.0	453	24	AAD32054	Pigmy rattlesnake
6	119	70.0	583	24	AAI42016	Snake adder snake
7	119	70.0	601	22	AAI71876	Snake venom blood
8	115	67.6	456	24	AAD32056	Pigmy rattlesnake
9	115	67.6	544	25	ABX93668	CDNA encoding Dein
10	111	65.3	690	16	AAQ89309	Snake venom antith
11	111	65.3	690	21	AAK61144	DNA encoding a sna
12	111	65.3	690	24	AAK99834	DNA encoding the a
13	110	64.7	580	24	AAD32057	Pigmy rattlesnake
14	103	60.6	432	24	AAD32058	Pigmy rattlesnake
15	99	58.2	632	24	AAI42015	Korean adder snake
16	97	57.1	725	24	AAD32059	Pigmy rattlesnake
17	94	55.3	716	17	AAI64829	Snake venom protei
18	91	53.5	474	24	AAD32060	Pigmy rattlesnake
19	84	49.4	492	22	AAI64154	Human zlecl encodi
20	84	49.4	492	24	AAD27975	Human lectin zlecl
21	84	49.4	5191	15	AAQ57710	Neurocan DNA. Hom
22	82	48.2	495	24	ABK32999	DNA encoding novel
23	82	48.2	629	24	AAI64153	Human lectin zlecl
24	82	48.2	629	24	AAD27974	Human lectin zlecl
25	82	48.2	655	22	AAK84316	Human EXCS encodin
26	82	48.2	2652	25	ABZ81727	Rat mutant brain-e
27	82	48.2	2652	25	ABZ81728	Rat brain-enriched
28	82	48.2	3153	22	AAQ29464	Murine brevidin cd
29	80	47.1	790	11	AAQ05300	Bovine pancreatic
30	79	46.5	3259	16	AAT05627	Bovine brevidin co
31	76	44.7	634	21	AAK69541	Human secreted pro
32	76	44.7	2878	25	AAI48133	Human chondroitin
33	76	44.7	3476	22	AAI91017	Human secreted pro
34	76	44.7	3476	24	ABK69986	CDNA encoding huma
35	75	44.1	2087	24	ABK62593	Rat sequence diffe
36	75	44.1	2465	24	ABK62593	Rat sequence diffe
37	75	44.1	4302	21	AAI77694	Human cancer assoc
38	75	44.1	4302	25	ABX63334	Human CDNA #334 di
39	75	44.1	8224	12	AAI12261	Versican gene. Ho
40	75	44.1	8224	24	AAI11088	Human breast cance
41	75	44.1	8224	25	ACC50121	Breast cancer asso
42	75	44.1	11185	24	ABN96814	Gene #3312 used to
43	75	44.1	11185	24	ABL62702	Colon adenocarcino
44	75	44.1	12319	24	AAI94985	Human DNA sequence
45	74	43.5	857	22	AAI31230	Human CDNA encodin

ALIGNMENTS

RESULT 1
AAD32053
ID AAD32053 standard; DNA; 483 BP.
XX
AC AAD32053;
XX
DT 18-JUN-2002 (first entry)
XX
DE Pigmy rattlesnake venom gland protein, Zsnk2 gene.
XX
KW Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;
platelet aggregation; gene; Zsnk2; ds.
XX
XX Sistrurus miliarius.
XX
XX Key Location/Qualifiers
FH 3..458
FT CDS

CC angiopathic thrombosis, cerebral thrombosis, ischaemic cerebral vascular
 CC diseases, unstable angina, acute thrombosis; unstable stenocardia,
 CC thromboangiitis obliterans, pulmonary embolism, deep vein thrombosis,
 CC peripheral arterial occlusion, stroke. It is also useful for treating
 CC atherosclerosis, oedema and inflammation, cancer and neurodegenerative
 CC diseases. The present sequence represents cDNA encoding the
 CC Deinagkistrodon acutus antithrombosis enzyme beta chain.
 XX

SQ Sequence 454 BP; 114 A; 107 C; 114 G; 109 T; 10 other;

Alignment Scores:

Pred. No.: 2.67e-10 Length: 454
 Score: 132.00 Matches: 21
 Percent Similarity: 79.31% Conservative: 2
 Best Local Similarity: 72.41% Mismatches: 6
 Query Match: 77.65% Indels: 0
 DB: 25 Gaps: 0

US-09-938-114-3 (1-29) x ABX93674 (1-454)

Qy 1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
 Db 72 GATTGTCCCTCTGATGGTCTCTCTATGAGGGCATTGCTACAGCCCTTCGATGAACCT 131

Qy 21 LysThrTrpThrAspAlaGluSerPhe 29

Db 132 AAGACCTGGGCAGATGCAGAGAAATTC 158

RESULT 3

AAI71877

ID AAI71877 standard; DNA; 633 BP.

AC AAI71877;

DT 10-JAN-2002 (first entry)

XX Snake venom blood anticoagulant halyxin B chain coding sequence.

XX Snake; pit viper; venom; halyxin; anticoagulant; thrombogenesis;
 thrombosis; ds.

XX Agkistrodon halys.

XX KR2001049671-A.

XX 15-JUN-2001.

XX 29-JUN-2000; 2000KR-0036591.

XX 29-JUN-1999; 99KR-0025105.

XX (BIOB-) BIOBUD CO LTD.

XX Jang YS, Jung GH, Kim DS, Koo BH, Son YD;

XX WPI; 2001-637330/73.

XX P-PSDB; AAM51544.

XX Halyxin as blood anticoagulation protein separated from snake venom -

XX Claim 1; Page 11; 21pp; Korean.

XX The invention relates to halyxin, a novel protein with very strong
 CC blood anticoagulation activity. The protein was separated from snake
 CC venom of Agkistrodon halys brevicaudus (a Korean pit viper) and can
 CC be used in the treatment of thrombogenesis. The present sequence
 CC encodes the B chain of halyxin.

SQ Sequence 633 BP; 179 A; 142 C; 157 G; 155 T; 0 other;

Alignment Scores:

Pred. No.: 2.19e-09 Length: 633
 Score: 127.00 Matches: 20

Percent Similarity: 75.86% Conservative: 2
 Best Local Similarity: 68.97% Mismatches: 7
 Query Match: 74.71% Indels: 0
 DB: 22 Gaps: 0

US-09-938-114-3 (1-29) x AAI71877 (1-633)

Qy 1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
 Db 93 GATTGTCCCTCTGATGGTCTCTCTATGAGGGCATTGCTACAAACCAATTTAATGAACA 152

Qy 21 LysThrTrpThrAspAlaGluSerPhe 29

Db 153 AAAAATTGGCCGATGCAGAGAAATTC 179

RESULT 4

AAD32055

XX AAD32055 standard; DNA; 721 BP.

AC AAD32055;

DT 18-JUN-2002 (first entry)

DE Pigmy rattlesnake venom gland protein, Zsnk3 gene.

XX Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;
 platelet aggregation; gene; Zsnk3; ds.

XX Sistrurus miliarius.

XX Key Location/Qualifiers

FT CDS 91..549

FT /*tag= a

FT /product= "Zsnk3 protein"

FT sig_peptide 91..159

FT /*tag= b

FT mat_peptide 160..546

FT /*tag= c

FT /product= "Mature Zsnk3 protein"

XX WO200214364-A2.

XX 21-FEB-2002.

XX 13-AUG-2001; 2001WO-US25310.

XX 14-AUG-2000; 2000US-225072P.

XX 14-AUG-2000; 2000US-225087P.

XX 15-AUG-2000; 2000US-225489P.

XX 15-AUG-2000; 2000US-225490P.

XX 20-DEC-2000; 2000US-356997P.

XX (ZYMO) ZYMOGENETICS INC.

XX Sheppard PO, Bishop PD;

XX WPI; 2002-269180/31.

XX P-PSDB; AAE20179.

XX New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which
 PT affect blood coagulation and platelet aggregation system, useful in
 PT therapy and diagnostics, or as tools in the study of genetics or
 PT molecular biology -

XX Claim 5; Page 73-74; 79pp; English.

XX The invention relates to new pigmy rattlesnake (Sistrurus miliarius)
 CC venom gland proteins, which affect blood coagulation and platelet
 CC aggregation system. The polypeptides, which affect blood coagulation and
 CC platelet aggregation system, are useful in therapy and diagnostics. The
 CC polypeptides are also useful as an educational tool in laboratory
 CC practical kits for courses related to genetics and molecular biology.
 CC protein chemistry and antibody production and analysis. The

CC polynucleotide or polypeptide can be used as standards or as unknowns
 CC for testing purposes. The polypeptides are also useful in identifying
 CC proteins by western blotting, protein purification, determining the
 CC weight of expressed polypeptides as a ratio to total protein expressed,
 CC identifying peptide cleavage sites, coupling amino and carboxyl terminal
 CC tags, mass spectrometry, circular dichroism to determine conformation or
 CC affinity chromatography columns to purify the protein, cloning or
 CC sequencing. The present sequence is *Sistrurus miliarius* venom gland
 CC protein, Zsnk3 gene.

XX Sequence 721 BP; 192 A; 167 C; 203 G; 159 T; 0 other;

Alignment Scores: 2.56e-09 Length: 721
 Pred. No.: 127.00 Matches: 21
 Score: 127.00
 Percent Similarity: 79.31% Conservative: 2
 Best Local Similarity: 72.41% Mismatches: 6
 Query Match: 74.71% Indels: 0
 DB: 24 Gaps: 0

US-09-938-114-3 (1-29) x AAD32055 (1-721)

Qy 1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
 Db 160 GATTGTCCCTCTGTTGGTCTCTATGATCAGCATTCGACAGGGCTTCAACAACATC 219
 Qy 21 LysThrTrpThrAspAlaGluSerPhe 29
 Db 220 AAGACGTGGGACGATGACAGAGGCTC 246

RESULT 5

AAD32054
 ID AAD32054 standard; DNA; 453 BP.

XX AAD32054;

DT 18-JUN-2002 (first entry)

XX Pigmy rattlesnake venom gland protein Zsnk2, degenerate nucleic acid.

DE Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;

KW platelet aggregation; ds.

XX *Sistrurus miliarius*.

XX WO200214364-A2.

XX 21-FEB-2002.

XX 13-AUG-2001; 2001WO-US25310.

XX 14-AUG-2000; 2000US-225072P.

XX 14-AUG-2000; 2000US-225087P.

XX 15-AUG-2000; 2000US-225489P.

XX 15-AUG-2000; 2000US-225490P.

XX 20-DEC-2000; 2000US-356997P.

XX (ZYMO) ZYMOGENETICS INC.

XX Sheppard PO, Bishop PD;

XX WPI; 2002-269180/31.

XX New pigmy rattlesnake (*Sistrurus miliarius*) venom gland proteins, which

XX affect blood coagulation and platelet aggregation system, useful in

XX therapy and diagnostics, or as tools in the study of genetics or

XX molecular biology

XX Disclosure; Page 72-73; 79pp; English.

XX The invention relates to new pigmy rattlesnake (*Sistrurus miliarius*)

XX venom gland proteins, which affect blood coagulation and platelet

XX aggregation system. The polypeptides, which affect blood coagulation and

CC platelet aggregation system, are useful in therapy and diagnostics. The
 CC polypeptides are also useful as an educational tool in laboratory
 CC practical kits for courses related to genetics and molecular biology,
 CC protein chemistry and antibody production and analysis. The
 CC polynucleotide or polypeptide can be used as standards or as unknowns
 CC for testing purposes. The polypeptides are also useful in identifying
 CC proteins by western blotting, protein purification, determining the
 CC weight of expressed polypeptides as a ratio to total protein expressed,
 CC identifying peptide cleavage sites, coupling amino and carboxyl terminal
 CC tags, mass spectrometry, circular dichroism to determine conformation or
 CC affinity chromatography columns to purify the protein, cloning or
 CC sequencing. The present sequence is *Sistrurus miliarius* venom gland
 CC protein Zsnk2, degenerate nucleic acid.

XX Sequence 453 BP; 74 A; 32 C; 84 G; 70 T; 193 other;

Alignment Scores: 2.17e-08 Length: 453
 Pred. No.: 119.00 Matches: 18
 Score: 119.00
 Percent Similarity: 68.97% Conservative: 2
 Best Local Similarity: 62.07% Mismatches: 9
 Query Match: 70.00% Indels: 0
 DB: 24 Gaps: 0

US-09-938-114-3 (1-29) x AAD32054 (1-453)

Qy 1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
 Db 58 GATTGYCCNWSNGAYTGGWSNWSNTAYGACRCATGYTAYARGTNTTWSNGARYTN 117

Qy 21 LysThrTrpThrAspAlaGluSerPhe 29

Db 118 AARACNTGGGAYGAYGCGARGWSNTTY 144

RESULT 6

AAL42016
 ID AAL42016 standard; cDNA; 583 BP.

XX AAL42016;

DT 16-MAY-2002 (first entry)

XX Korean adder snake venom salmorin B chain protein cDNA sequence.

DE Korean adder; ss; salmorin protein B chain; snake venom;

KW fibrinogen clotting inhibition; thrombosis; prothrombin binding;

KW thrombin binding; blood coagulation.

XX Agkistrodon halys brevicaudus.

XX Key Location/Qualifiers

FT CDS 7..444

FT /*tag= a "salmorin B chain protein"

FT /*product= 7..75

FT sig_peptide /*tag= b

FT /*note= "Signal peptide"

FT mat_peptide 76..441

FT /*tag= c

FT /*note= "Mature Salmorin B chain protein"

FT 3'UTR 445..570

FT /*tag= d

FT polyA_signal 571..583

FT /*tag= e

XX WO200214514-A1.

XX 21-FEB-2002.

XX 26-JUL-2001; 2001WO-KR01277.

XX 27-JUL-2000; 2000KR-0043470.

PA (BIOB-) BIOBUD CO LTD.
 XX
 PI Chung K, Kim D, Koh Y;
 XX
 DR WPI; 2002-241907/29.
 DR P-PSDB; AAO14521.
 XX
 PT New salmorin protein derived from venom of Korean adder Agkistrodon
 PT halyx brevicaudus, useful for treating thrombosis by repressing
 PT fibrinogen clotting through repression of activation of prothrombin
 PT into thrombin
 XX
 PS Claim 1; Fig 1B; 30pp; English.
 XX
 CC The invention comprises the nucleotide and protein sequences of a
 CC salmorin protein derived from the venom of Korean adder. Salmorin protein
 CC is composed of an A chain and a B chain, and has inhibitory activity
 CC against fibrinogen clotting. The salmorin protein of the invention is
 CC useful for treating thrombosis, as it represses fibrinogen clotting
 CC potentially by binding to prothrombin and thrombin so as to delay blood
 CC coagulation. The present cDNA sequence encodes the Korean adder salmorin
 CC B chain protein.
 XX
 SQ Sequence 583 BP; 151 A; 137 C; 150 G; 145 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 2.96e-08 Length: 583
 Score: 119.00 Matches: 19
 Percent Similarity: 68.97% Conservative: 1
 Best Local Similarity: 65.52% Mismatches: 9
 Query Match: 70.00% Indels: 0
 DB: 24 Gaps: 0
 US-09-938-114-3 (1-29) x AAL42016 (1-583)
 QY 1 AspCysSerSerAspTrpSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
 DB 76 GATGTCCCTCTGGTGGTCTCTATGAGGGCATTGCTACAGCTCTTCACATCAACAG 135
 QY 21 LysThrTrpThrAspAlaGluSerPhe 29
 DB 136 TCTCAGTGGGCCCATGCGAGAAATTC 162
 RESULT 7
 AAI71876
 ID AAI71876 standard; DNA; 601 BP.
 AC AAI71876;
 XX
 XX 10-JAN-2002 (first entry)
 DE Snake venom blood anticoagulant halyxin A chain coding sequence.
 DE
 DE Snake; pit viper; venom; halyxin; anticoagulant; thrombogenesis;
 KW thrombosis; ds.
 KW
 XX Agkistrodon halyx.
 OS
 XX KR2001049671-A.
 PN
 XX 15-JUN-2001.
 PD
 XX 29-JUN-2000; 2000KR-0036591.
 PF
 XX 29-JUN-1999; 99KR-0025105.
 PR
 XX (BIOB-) BIOBUD CO LTD.
 PA
 XX Jang YS, Jung GH, Kim DS, Koo BH, Son YD;
 PI WPI; 2001-637330/73.
 DR P-PSDB; AAM51543.
 XX

PT Halyxin as blood anticoagulation protein separated from snake venom -
 XX Claim 1; Page 9; 21pp; Korean.
 PS
 XX The invention relates to halyxin, a novel protein with very strong
 CC blood anticoagulation activity. The protein was separated from snake
 CC venom of Agkistrodon halyx brevicaudus (a Korean pit viper) and can
 CC be used in the treatment of thrombogenesis. The present sequence
 CC encodes the A chain of halyxin.
 XX
 SQ Sequence 601 BP; 166 A; 133 C; 163 G; 139 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 3.08e-08 Length: 601
 Score: 119.00 Matches: 19
 Percent Similarity: 72.41% Conservative: 2
 Best Local Similarity: 65.52% Mismatches: 8
 Query Match: 70.00% Indels: 0
 DB: 22 Gaps: 0
 US-09-938-114-3 (1-29) x AAI71876 (1-601)
 QY 1 AspCysSerSerAspTrpSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
 DB 93 GATGTCCCTCTGGTGGTCTCTATGAGGGCATTGCTACAGCTCTTCATCTTC 152
 QY 21 LysThrTrpThrAspAlaGluSerPhe 29
 DB 153 AAGACCTGGCGCAGAGCAGAGAGGTTTC 179
 RESULT 8
 AAD32056
 ID AAD32056 standard; DNA; 456 BP.
 XX
 AC AAD32056;
 XX
 XX 18-JUN-2002 (first entry)
 DT
 XX Pigmy rattlesnake venom gland protein Zsnk3, degenerate nucleic acid.
 DE
 DE Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;
 KW platelet aggregation; ds.
 KW
 XX Sistrurus miliarius.
 OS
 XX WO200214364-A2.
 PN
 XX 21-FEB-2002.
 PD
 XX 13-AUG-2001; 2001WO-US25310.
 PF
 XX 14-AUG-2000; 2000US-225072P.
 PR 14-AUG-2000; 2000US-225087P.
 PR 15-AUG-2000; 2000US-225489P.
 PR 15-AUG-2000; 2000US-225490P.
 PR 20-DEC-2000; 2000US-356997P.
 XX
 XX (ZYMO) ZYMOGENETICS INC.
 PA
 XX Sheppard PO, Bishop PD;
 PI
 XX WPI; 2002-269180/31.
 DR P-PSDB; AAE20179.
 DR
 XX New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which
 PT affect blood coagulation and platelet aggregation system, useful in
 PT therapy and diagnostics, or as tools in the study of genetics or
 PT molecular biology
 PT
 XX Disclosure; Page 74; 79pp; English.
 PS
 XX The invention relates to new pigmy rattlesnake (Sistrurus miliarius)
 CC venom gland proteins, which affect blood coagulation and platelet
 CC

CC aggregation system. The polypeptides, which affect blood coagulation and
 CC platelet aggregation system, are useful in therapy and diagnostics. The
 CC polypeptides are also useful as an educational tool in laboratory
 CC practical kits for courses related to genetics and molecular biology,
 CC protein chemistry and antibody production and analysis. The
 CC polynucleotide or polypeptide can be used as standards or as unknowns
 CC for testing purposes. The polypeptides are also useful in identifying
 CC proteins by western blotting, protein purification, determining the
 CC weight of expressed polypeptides as a ratio to total protein expressed,
 CC identifying peptide cleavage sites, coupling amino and carboxyl terminal
 CC tags, mass spectrometry, circular dichroism to determine conformation or
 CC affinity chromatography columns to purify the protein, cloning or
 CC sequencing. The present sequence is *Sistrurus miliarius* venom gland
 CC protein Zsnk3, degenerate nucleic acid.

SQ Sequence 456 BP; 72 A; 31 C; 92 G; 66 T; 195 other;

Alignment Scores:
 Pred. No.: 8.49e-08 Length: 456
 Score: 115.00 Matches: 18
 Percent Similarity: 65.52% Conservative: 1
 Best Local Similarity: 62.07% Mismatches: 10
 Query Match: 67.65% Indels: 0
 DB: 24 Gaps: 0

US-09-938-114-3 (1-29) x AAB32056 (1-456)

QY 1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
 DB 70 GAYTGYCCNWSNGCTGGWSNWSNTAYGAYCARCAITGYTAYMGNGINTTVAARCAVTN 129

QY 21 LysThrTrpThrAspAlaGluSerPhe 29

DB 130 AARACNTGGGAYGAYGCGNARMGNTTY 156

RESULT 9

ABX93668
 ID ABX93668 standard; cDNA; 544 BP.

AC ABX93668;

DT 02-JUN-2003 (first entry)

XX cDNA encoding Deinagkistrodon acutus antithrombosis enzyme alpha chain.

XX Antithrombosis; ss; PCR; alpha chain; fibrin hydrolysis; blood clot;
 KW platelet aggregation; vaso-occlusive disorder; thromboembolic disease;
 KW myocardial infarction; restenosis; cancer; neurodegenerative disease;
 KW angioathic thrombosis; cerebral thrombosis; thromboangiitis obliterans;
 KW ischaemic cerebral vascular disease; unstable angina; acute thrombosis;
 KW unstable stenocardia; pulmonary embolism; deep vein thrombosis; oedema;
 KW peripheral arterial occlusion; stroke; atherosclerosis; inflammation;
 XX thrombosis.

OS Deinagkistrodon acutus.

XX Key Location/Qualifiers

FT CDS 1..363
 FT /tag= a
 FT /product= "Antithrombosis enzyme alpha chain"
 FT /partial
 FT /transl_except= (pos:244..246,aa:Lys)
 FT /note= "No start codon given"

XX US6489451-B1.

PN 03-DEC-2002.

XX 10-APR-1998; 98US-0058740.

XX 10-APR-1997; 97US-043886P.

PA (HEFE-) HEFEI SIU FUNG USTC PHARM CO LTD.

XX

PI Li BX, Cheng X;

XX WPI; 2003-352116/33.

DR P-PSDB; ABU08798.

XX New purified Agkistrodon actus anti-thrombosis enzyme, useful for
 PT preventing and treating vaso-occlusive and thromboembolic disorders,
 PT including myocardial infarction, restenosis, cerebral thrombosis and
 PT unstable angina -

XX Example 3; Column 25-26; 19pp; English.

XX The invention relates to a new Deinagkistrodon actus anti-thrombosis
 CC enzyme, where the enzyme hydrolyses fibrin, dissolves blood clots and
 CC prevents platelet aggregation. The anti-thrombosis enzyme was
 CC administered to rabbits intravenously. Thrombosis was determined before
 CC and after the administration at 0.5, 1.0, 2.0 and 3.0 hours by the
 CC Chandler method. The enzyme showed anti-thrombosis activity at 0.5 hour
 CC following administration at 0.005 micro/kg and this activity was
 CC increased significantly at 1.0 hour and at 0.01 micro /kg. The
 CC composition is useful for preventing and treating vaso-occlusive and
 CC thromboembolic disorders, including myocardial infarction, restenosis,
 CC angioathic thrombosis, cerebral thrombosis, ischaemic cerebral vascular
 CC diseases, unstable angina, acute thrombosis, unstable stenocardia,
 CC thromboangiitis obliterans, pulmonary embolism, deep vein thrombosis,
 CC peripheral arterial occlusion, stroke. It is also useful for treating
 CC atherosclerosis, oedema and inflammation, cancer and neurodegenerative
 CC diseases. The present sequence represents cDNA encoding the
 CC Deinagkistrodon acutus antithrombosis enzyme alpha chain.

SQ Sequence 544 BP; 158 A; 122 C; 143 G; 121 T; 0 other;

Alignment Scores:
 Pred. NO.: 1.05e-07 Length: 544
 Score: 115.00 Matches: 20
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 67.65% Indels: 0
 DB: 25 Gaps: 0

US-09-938-114-3 (1-29) x ABX93668 (1-544)

QY 10 GluGlyHisCysTyrLysValPheLysGlnSerLysThrTrpThrAspAlaGluSerPhe 29
 DB 1 GAAGGGCATTCCTACAGGCTTCAACACATCTAGACCTGGACAGATGCGAGAGCTTC 60

RESULT 10

AAQ89309
 ID AAQ89309 standard; cDNA; 690 BP.

AC AAQ89309;

XX 25-MAR-2003 (updated)

DT 28-NOV-1995 (first entry)

XX Snake venom antithrombotic oligopeptide cDNA.

XX Antithrombotic peptide; snake venom; platelet binding inhibition;
 KW von Willebrand factors; Crotalus horridus horridus; db.

OS Crotalus horridus horridus.

XX Key Location/Qualifiers

FT CDS 66..515
 FT /tag= a

XX WO9508573-A1.

XX 30-MAR-1995.

XX 21-SEP-1994; 94WO-JP01555.

XX

```
PR 22-SEP-1993; 93JP-0236975.
XX (AJIN ) AJINOMOTO KK.
PA Fukuchi N, Ishii K, Kito M, Kobayashi T, Nagano M;
PI Tanaka A, Yamamoto H, Yoshimoto R;
XX WPI; 1995-139559/18.
DR P-PSDB; AAR71981.
XX Single-chain antithrombotic peptide - obtained by cleaving an
PT oligopeptide from snake venom to break inter-chain di:sulphide
PT bonds but preserve intra-chain di:sulphide bonds
XX Example 2; Pages 47-48; 84pp; Japanese.
XX AAR71978 and AAR71979 are snake venom derived antithrombotic peptides,
CC specifically from the snake venom oligopeptide AAR71981, encoded by
CC AAR71979. These peptides have the advantage of avoiding significant
CC thrombocytopenia when administered at the minimum dose, for in vivo
CC inhibition of platelet von Willebrand factor binding.
CC (Updated on 25-MAR-2003 to correct FN field.)
XX SQ Sequence 690 BP; 167 A; 180 C; 170 G; 173 T; 0 other;

Alignment Scores:
Pred. No.: 5.47e-07 Length: 690
Score: 111.00 Matches: 18
Percent Similarity: 72.41% Conservative: 3
Best Local Similarity: 62.07% Mismatches: 8
Query Match: 65.29% Indels: 0
DB: 16 Gaps: 0

US-09-938-114-3 (1-29) x AAR71979 (1-690)
QY 1 AspCysSerSerAspTrpSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
Db 141 GAATGTCCTCCGGTGGTCTCTCTATGATCGGTATGCTACAGCCCTTCACACAGAG 200
QY 21 LysThrTrpThrAspAlaGluSerPhe 29
Db 201 ATGACCTGGGCCGATGCAGAGAGGTTTC 227

RESULT 11
AAC611144
ID AAC611144 standard; DNA; 690 BP.
XX AC AAC611144;
XX 07-FEB-2001 (first entry)
XX DNA encoding a snake venom derived protein.
XX Subunit peptide production; snake venom; rattlesnake; thrombolytic;
XX von Willebrand's factor; blood platelet-inhibitory activity; ds.
XX Crotalus horridus horridus.
XX WO200059926-A1.
XX 12-OCT-2000.
XX 31-MAR-2000; 2000MO-JF02127.
XX 02-APR-1999; 99JP-0096073.
XX (AJIN ) AJINOMOTO CO INC.
XX Fukuchi N, Kageyama S, Kito M, Kayahara T, Yamamoto H;
XX WPI; 2000-664985/64.
XX P-PSDB; AAY85628.
XX

PT Producing physiologically-active subunit peptides originating in
PT polymer proteins by denaturation and specific separation, with lower
PT antigenicity but improved solubility and stability, e.g. blood
XX platelet-binding inhibitors -
XX Disclosure; Page 46; 51pp; Japanese.
XX This invention relates to a method for the production of a subunit
CC peptide originating from a polymer protein with disulphide bonds within
CC and between subunits. The method comprises denaturing the protein or its
CC subunit using a protein denaturing agent in a solution, removing the
CC agent in the presence of a polyoxyalkyl polyether which reacts with a
CC thiol group and unwinds the subunit, and separating the polyoxyalkyl
CC polyether-bound subunit peptide. The method can be used for producing
CC physiologically-active subunit peptides for polymer proteins e.g. snake
CC venom-originated dimer peptide with blood platelet-inhibitory activity
CC on von Willebrand's factor. The peptides produced have platelet-binding
CC inhibitory, and thrombolytic activity. The present sequence represents
CC DNA encoding rattlesnake protein used in an example illustrating the
XX method of the invention.
XX SQ Sequence 690 BP; 167 A; 180 C; 170 G; 173 T; 0 other;

Alignment Scores:
Pred. No.: 5.47e-07 Length: 690
Score: 111.00 Matches: 18
Percent Similarity: 72.41% Conservative: 3
Best Local Similarity: 62.07% Mismatches: 8
Query Match: 65.29% Indels: 0
DB: 21 Gaps: 0

US-09-938-114-3 (1-29) x AAC61144 (1-690)
QY 1 AspCysSerSerAspTrpSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
Db 141 GAATGTCCTCCGGTGGTCTCTCTATGATCGGTATGCTACAGCCCTTCACACAGAG 200
QY 21 LysThrTrpThrAspAlaGluSerPhe 29
Db 201 ATGACCTGGGCCGATGCAGAGAGGTTTC 227

RESULT 12
AAC99834
ID AAC99834 standard; DNA; 690 BP.
XX AC AAC99834;
XX 19-JUL-2002 (first entry)
XX DNA encoding the antithrombotic wild-type rattlesnake protein.
XX Thrombolytic; glycoprotein Ib-binding protein; snake venom; rattlesnake;
XX antithrombotic; glycoprotein Ib; long half life; low antigenicity;
XX drug efficacy; gene; ds.
XX Crotalus horridus horridus.
XX Key Location/Qualifiers
XX CDS 66..515
XX /*tag= a
XX /product= "Antithrombotic wild-type rattlesnake protein"
XX BP1195384-A1.
XX 10-APR-2002.
XX 04-OCT-2001; 2001EP-0123277.
XX 04-OCT-2000; 2000JP-0305279.
XX (AJIN ) AJINOMOTO CO INC.
XX Fukuchi N, Kito M, Kayahara T, Futaki F, Ishikawa K, Suzuki E;
PI
```

```

PI Gondoh K, Shimba N, Yamada N;
XX WPI; 2002-364482/40.
DR P-PSDB; AAO20974.
XX Glycoprotein Ib-binding protein, useful for treating thrombosis,
PT comprises specific mutations in protein originating from snake venom -
PS Example 1; Page 25; 49pp; English.
XX
XX The invention relates to a glycoprotein Ib-binding protein, originating
CC from snake venom, comprising specific mutations and antithrombotic
CC activity. Glycoprotein Ib-binding protein is used in a drug having
CC antithrombotic activity. Glycoprotein Ib-binding protein has a high
CC binding activity to glycoprotein Ib, a long half life/drug efficacy
CC retention in blood, and low antigenicity. This polynucleotide sequence
CC represents DNA of the wild-type rattlesnake protein of the invention.
XX
XX Sequence 690 BP; 167 A; 180 C; 170 G; 173 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 5.47e-07 Length: 690
Score: 111.00 Matches: 18
Percent Similarity: 72.41% Conservative: 3
Best Local Similarity: 62.07% Mismatches: 8
Query Match: 65.29% Indels: 0
DB: 24 Gaps: 0
US-09-938-114-3 (1-29) x AAK99834 (1-690)
Qy 1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
Db 141 GAATGTCCTCCGGTGGTCTCTCTATGATCGGTATGCTACAGGCCCTTCACACAGAG 200
Qy 21 LysThrTrpThrAspAlaGluSerPhe 29
Db 201 ATGACCTGGGCGGATGCAGAGAGGTTTC 227
RESULT 13
AAD32057
ID AAD32057 standard; DNA; 580 BP.
AC AAD32057;
XX 18-JUN-2002 (first entry)
XX Pigmy rattlesnake venom gland protein, Zsnk4 gene.
XX Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;
XX platelet aggregation; gene; Zsnk4; ds.
XX Sistrurus miliarius.
XX Key Location/Qualifiers
XX CDS 3..437
XX /*tag= a
XX /product= "Zsnk4 protein"
XX /note= "CDS does not include start codon"
XX /partial
XX sig_peptide 3..29
XX /*tag= b
XX mat_peptide 30..434
XX /*tag= c
XX /product= "Mature Zsnk4 protein"
XX
XX WO200214364-A2.
XX 21-FEB-2002.
XX 13-AUG-2001; 2001WO-US25310.
XX 14-AUG-2000; 2000US-225072P.
XX 14-AUG-2000; 2000US-225087P.

```

```

PR 15-AUG-2000; 2000US-225489P.
PR 15-AUG-2000; 2000US-225490P.
PR 20-DEC-2000; 2000US-356997P.
XX (ZYMO ) ZYMOGENETICS INC.
XX Sheppard PO, Bishop PD;
XX WPI; 2002-269180/31.
DR P-PSDB; AAE20180.
XX New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which
PT affect blood coagulation and platelet aggregation system, useful in
PT therapy and diagnostics, or as tools in the study of genetics or
XX molecular biology -
XX Claim 5; Page 75; 79pp; English.
XX The invention relates to new pigmy rattlesnake (Sistrurus miliarius)
CC venom gland proteins, which affect blood coagulation and platelet
CC aggregation system. The polypeptides, which affect blood coagulation and
CC platelet aggregation system, are useful in therapy and diagnostics. The
CC polypeptides are also useful as an educational tool in laboratory
CC practical kits for courses related to genetics and molecular biology,
CC protein chemistry and antibody production and analysis. The
CC polynucleotide or polypeptide can be used as standards or as unknowns
CC for testing purposes. The polypeptides are also useful in identifying
CC proteins by western blotting, protein purification, determining the
CC weight of expressed polypeptides as a ratio to total protein expressed,
CC identifying peptide cleavage sites, coupling amino and carboxyl terminal
CC tags, mass spectrometry, circular dichroism to determine conformation or
CC affinity chromatography columns to purify the protein, cloning or
CC sequencing. The present sequence is Sistrurus miliarius venom gland
CC protein, Zsnk4 gene.
XX
XX Sequence 580 BP; 154 A; 133 C; 163 G; 130 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 6.2e-07 Length: 580
Score: 110.00 Matches: 17
Percent Similarity: 75.86% Conservative: 5
Best Local Similarity: 58.62% Mismatches: 7
Query Match: 64.71% Indels: 0
DB: 24 Gaps: 0
US-09-938-114-3 (1-29) x AAD32057 (1-580)
Qy 1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
Db 36 GATTGTCCCTCTGATTGGTATGCTATGATCAGTATTGTCACAGGGTCATCAACAACTC 95
Qy 21 LysThrTrpThrAspAlaGluSerPhe 29
Db 96 AGGACGTGGGACGATGCAGAGAGGTTTC 122
RESULT 14
AAD32058
ID AAD32058 standard; DNA; 432 BP.
XX
XX AAD32058;
XX 18-JUN-2002 (first entry)
XX Pigmy rattlesnake venom gland protein Zsnk4, degenerate nucleic acid.
XX Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;
XX platelet aggregation; ds.
XX Sistrurus miliarius.
XX WO200214364-A2.
XX 21-FEB-2002.

```

```

XX PF 13-AUG-2001; 2001WO-US25310.
XX PR 14-AUG-2000; 2000US-225072P.
XX PR 14-AUG-2000; 2000US-225087P.
XX PR 15-AUG-2000; 2000US-225489P.
XX PR 15-AUG-2000; 2000US-225490P.
XX PR 20-DEC-2000; 2000US-356997P.
XX PA (ZYMO ) ZYMOGENETICS INC.
XX XX Sheppard PO, Bishop PD;
XX XX MPI; 2002-269180/31.
XX XX
XX XX New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which
XX FT affect blood coagulation and platelet aggregation system, useful in
XX FT therapy and diagnostics, or as tools in the study of genetics or
XX FT molecular biology
XX XX
XX PS Disclosure; Page 76; 79pp; English.
XX XX
XX CC The invention relates to new pigmy rattlesnake (Sistrurus miliarius)
XX CC venom gland proteins, which affect blood coagulation and platelet
XX CC aggregation system. The polypeptides, which affect blood coagulation and
XX CC platelet aggregation system, are useful in therapy and diagnostics. The
XX CC polypeptides are also useful as an educational tool in laboratory
XX CC practical kits for courses related to genetics and molecular biology,
XX CC protein chemistry and antibody production and analysis. The
XX CC polynucleotide or polypeptide can be used as standards or as unknowns
XX CC for testing purposes. The polypeptides are also useful in identifying
XX CC proteins by western blotting, protein purification, determining the
XX CC weight of expressed polypeptides as a ratio to total protein expressed,
XX CC identifying peptide cleavage sites, coupling amino and carboxyl terminal
XX CC tags, mass spectrometry, circular dichroism to determine conformation or
XX CC affinity chromatography columns to purify the protein, cloning or
XX CC sequencing. The present sequence is Sistrurus miliarius venom gland
XX CC protein Zsnk4, degenerate nucleic acid.
XX XX
XX SQ Sequence 432 BP; 91 A; 34 C; 84 G; 55 T; 168 other;

Alignment Scores:
Pred. No.: 4.62e-06 Length: 432
Score: 103.00 Matches: 16
Percent Similarity: 65.52% Conservative: 3
Best Local Similarity: 55.17% Mismatches: 10
Query Match: 60.59% Indels: 0
DB: 24 Gaps: 0

US-09-938-114-3 (1-29) x AAD32058 (1-432)

QY 1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
Db 34 GAYTGYCCNWSNGAYTGGTAYGCNTAYGTAYGNTAYMGNTNATHARCARNTN 93
QY 21 LysThrTrpThrAspAlaGluSerPhe 29
Db 94 MGNACNTGGGAYGAYCGNGARMGNTTY 120

RESULT 15
AAL42015
ID AAL42015 standard; cDNA; 632 BP.
XX AC AAL42015;
XX XX
XX DT 16-MAY-2002 (first entry)
XX XX Korean adder snake venom salmorin A chain protein cDNA sequence.
XX XX Korean adder; ss; salmorin protein A chain; snake venom;
XX KW fibrinogen clotting inhibition; thrombosis; prothrombin binding;
XX KW thrombin binding; blood coagulation.
XX XX

```

```

OS XX Agkistrodon halys brevicaudus.
XX FH Key Location/Qualifiers
XX FT CDS 34..498
XX FT /*tag= a
XX FT /product= "Salmorin A chain protein"
XX FT sig_peptide 34..102
XX FT /*tag= b
XX FT /note= "Signal peptide"
XX FT mat_peptide 103..495
XX FT /*tag= c
XX FT /note= "Mature Salmorin A chain protein"
XX FT 3'UTR 499..622
XX FT /*tag= d
XX FT polyA_signal 623..632
XX FT /*tag= e
XX PN WO200214514-A1.
XX XX 21-FEB-2002.
XX XX 26-JUL-2001; 2001WO-KR01277.
XX XX 27-JUL-2000; 2000KR-0043470.
XX XX (BIOB-) BIOBUD CO LTD.
XX XX Chung K, Kim D, Koh Y;
XX XX WPI; 2002-241907/29.
XX XX P-PSDB; AAO14520.
XX XX
XX FT New salmorin protein derived from venom of Korean adder Agkistrodon
XX FT halys brevicaudus, useful for treating thrombosis by repressing
XX FT fibrinogen clotting through repression of activation of prothrombin
XX FT into thrombin
XX XX
XX PS Claim 1; Fig 1A; 30pp; English.
XX XX
XX CC The invention comprises the nucleotide and protein sequences of a
XX CC salmorin protein derived from the venom of Korean adder. Salmorin protein
XX CC is composed of an A chain and a B chain, and has inhibitory activity
XX CC against fibrinogen clotting. The salmorin protein of the invention is
XX CC useful for treating thrombosis, as it represses fibrinogen clotting
XX CC potentially by binding to prothrombin and thrombin so as to delay blood
XX CC coagulation. The present cDNA sequence encodes the Korean adder salmorin
XX CC A chain protein.
XX XX
XX SQ Sequence 632 BP; 162 A; 140 C; 182 G; 148 T; 0 other;

Alignment Scores:
Pred. No.: 2.86e-05 Length: 632
Score: 99.00 Matches: 16
Percent Similarity: 60.71% Conservative: 1
Best Local Similarity: 57.14% Mismatches: 11
Query Match: 58.24% Indels: 0
DB: 24 Gaps: 0

US-09-938-114-3 (1-29) x AAL42015 (1-632)

QY 2 CysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSerLys 21
Db 112 TGTCCCTCTGGTTGGGGCAGCAATATGCTATGCTACCGCCCTTCAATCAACGATG 171
QY 22 ThrTrpThrAspAlaGluSerPhe 29
Db 172 ACCTGGGAAGATGCAGAGAGGTTTC 195

Search completed: December 8, 2003, 14:04:47
Job time : 100.594 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 8, 2003, 13:49:04 ; Search time 22.8686 Seconds
(without alignments)
559.725 Million cell updates/sec

Title: US-09-938-114-3

Perfect score: 170

Sequence: 1 DCSDSYSGHGVKVKQSKTWDAESF 29

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US0938114/runat_08122003_091001_22878/app_query.fasta_1.725
-DB-issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MAVRX=blosum62 -TRANS-human40.cdi
-LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0938114 @CGN 1 1 95 @runat_08122003_091001_22878 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.*

1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	132	77.6	454	4	US-09-058-740-12
2	115	67.6	544	4	US-09-058-740-1
3	111	65.3	690	2	US-08-612-840A-7
4	84	49.4	5191	1	US-08-340-428B-1
5	84	49.4	5191	5	PCT-US93-07306-1
6	79	46.5	3259	6	PCT-US95-03747-1
7	75	44.1	8224	6	5180808-1
8	73	42.9	402	3	US-08-543-246B-10
9	73	42.9	648	3	US-08-543-246B-14
10	73	42.9	1740	2	US-09-055-095-2
11	73	42.9	1755	3	US-08-543-246B-8
12	73	42.9	1841	4	US-09-996-243-318

13	73	42.9	4588	3	US-08-840-062-1	Sequence 1, Appli
14	73	42.9	4771	3	US-08-840-062-3	Sequence 3, Appli
15	72	42.4	4978	1	US-08-220-603A-1	Sequence 1, Appli
16	71	41.8	522	1	US-07-778-156-4	Sequence 4, Appli
17	71	41.8	522	2	US-08-422-166-4	Sequence 4, Appli
18	71	41.8	600	5	PCT-US93-10418-1	Sequence 1, Appli
19	71	41.8	602	3	US-09-385-982-257	Sequence 257, App
20	71	41.8	797	2	US-08-464-637-1	Sequence 1, Appli
21	71	41.8	797	2	US-08-822-261-5	Sequence 5, Appli
22	71	41.8	797	4	US-09-226-852-5	Sequence 5, Appli
23	71	41.8	798	1	US-07-778-156-12	Sequence 12, Appli
24	71	41.8	798	2	US-08-822-261-6	Sequence 6, Appli
25	71	41.8	798	2	US-08-422-166-12	Sequence 12, Appli
26	71	41.8	798	3	US-09-146-969-2	Sequence 2, Appli
27	71	41.8	798	4	US-09-226-852-6	Sequence 6, Appli
28	69	40.6	1104	3	US-09-111-470-1	Sequence 1, Appli
29	69	40.6	2076	4	US-09-489-847-51	Sequence 51, Appli
30	68	40.0	614	2	US-08-729-103-2	Sequence 2, Appli
31	68	40.0	1114	2	US-08-468-413-1	Sequence 1, Appli
32	68	40.0	1114	3	US-09-162-508-1	Sequence 1, Appli
33	68	40.0	1114	5	PCT-US95-07169-1	Sequence 1, Appli
34	67	39.4	43804	4	US-09-171-461-1	Sequence 1, Appli
35	66	38.8	360	3	US-08-543-246B-4	Sequence 4, Appli
36	66	38.8	405	3	US-08-543-246B-3	Sequence 3, Appli
37	66	38.8	405	3	US-08-543-246B-7	Sequence 7, Appli
38	66	38.8	645	3	US-08-543-246B-12	Sequence 12, Appli
39	66	38.8	693	3	US-08-543-246B-13	Sequence 13, Appli
40	66	38.8	699	3	US-08-543-246B-11	Sequence 11, Appli
41	66	38.8	1222	3	US-08-543-246B-5	Sequence 5, Appli
42	66	38.8	1223	4	US-09-016-434-1347	Sequence 1347, Ap
43	66	38.8	1333	3	US-08-543-246B-15	Sequence 15, Appli
44	66	38.8	1387	3	US-08-543-246B-1	Sequence 1, Appli
45	66	38.8	1643	4	US-09-517-605-16	Sequence 16, Appli

ALIGNMENTS

RESULT 1

US-09-058-740-12
; Sequence 12, Application US/09058740
; Patent No. 6489451
; GENERAL INFORMATION:

APPLICANT: Chun Wang, Benjamin XY Li, Xin Cheng,
Jing Liu, Li-Wen Niu, Wan-Zhi Huang,
Zhen-Yu Xu, Dan Luo, Lian-Di Kang,
Jin-Guo Ding, Fang Rong, Yan Liu and
Hui-Ran Chen
TITLE OF INVENTION: AN ANTI-THROMBOSIS ENZYME FROM THE SNAKE
VENOM OF AGKISTRODON ACUTUS

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon

STREET: 633 West Fifth Street

CITY: Suite 4700

STATE: California

COUNTRY: U.S.A.

ZIP: 90071-2066

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: FastSeq for Windows 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/058,740

FILING DATE: 10-Apr-1998

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: <Unknown>

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Chen, Anthony C.

REGISTRATION NUMBER: 38,673
REFERENCE/DOCKET NUMBER: 233/298
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 454 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 3...440
OTHER INFORMATION: "N" stands for any base.
"Xaa" stands for any amino acid.
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-058-740-12

Alignment Scores: Length: 454
Pred. No.: 1.73e-11 Matches: 21
Score: 132.00
Percent Similarity: 79.31% Conservative: 2
Best Local Similarity: 72.41% Mismatches: 6
Query Match: 77.65% Indels: 0
DB: 4 Gaps: 0

US-09-938-114-3 (1-29) x US-09-058-740-12 (1-454)

Qy 1 AspCysSerAspTyrSerSerTyrGluGlyHisCysTyrLysValPheGlnSer 20
Db 72 GATGTCTCTGAGTGTCTCTCTATGAGGGCATGCTACAGCCCTTCGATGAACCT 131
Qy 21 LysThrTrpThrAspAlaGluSerPhe 29
Db 132 AAGACCTGGGCAGATGCAGAGAAATTC 158

RESULT 2

US-09-058-740-1

; Sequence 1, Application US/09058740

; Patent No. 6489451

; GENERAL INFORMATION:

; APPLICANT: Chun Wang, Benjamin XY Li, Xin Cheng,
Jing Liu, Li-Wen Niu, Wan-Zhi Huang,
Zhen-yu Xu, Dan Luo, Lian-Di Kang,
Jin-Guo Ding, Fang Rong, Yan Liu and
Hui-Ran Chen

; TITLE OF INVENTION: AN ANTITHROMBOSIS ENZYME FROM THE SNAKE

; VENOM OF AGKISTRODON ACUTUS

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Lyon & Lyon

; STREET: 633 West Fifth Street

; Suite 4700

; CITY: Los Angeles

; STATE: California

; COUNTRY: U.S.A.

; ZIP: 90071-2066

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

; Storage

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: IBM P.C. DOS 5.0

; SOFTWARE: FastSeq for Windows 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/058,740

; FILING DATE: 10-Apr-1998

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: <Unknown>

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Chen, Anthony C.
; REGISTRATION NUMBER: 38,673
; REFERENCE/DOCKET NUMBER: 233/298
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 544 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-058-740-1

Alignment Scores: Length: 544
Pred. No.: 8.2e-09 Matches: 20
Score: 115.00
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 67.65% Indels: 0
DB: 4 Gaps: 0

US-09-938-114-3 (1-29) x US-09-058-740-1 (1-544)

Qy 10 GluGlyHisCysTyrLysValPheLysGlnSerLysThrTrpThrAspAlaGluSerPhe 29
Db 1 GAAGGGCATGCTACAGGCTCTCAACAAATCTAAGACCTGGACAGATGCAGAGACTTC 60

RESULT 3

US-08-612-840A-7

; Sequence 7, Application US/08612840A

; Patent No. 5856126

; GENERAL INFORMATION:

; APPLICANT: FUKUCHI, Naoyuki

; APPLICANT: YAMAMOTO, Hiroshi

; APPLICANT: NAGANO, Mitsuyo

; APPLICANT: KITO, Morikazu

; APPLICANT: TANAKA, Akiko

; APPLICANT: ISHII, Koichi

; APPLICANT: KOBAYASHI, Tsuyoshi

; APPLICANT: YOSHIMOTO, Ryota

; TITLE OF INVENTION: PEPTIDE HAVING ANTI-THROMBUS ACTIVITY AND

; METHOD OF PRODUCING THE SAME

; NUMBER OF SEQUENCES: 15

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Oblon, Spivak, McClelland, Maier & Neustadt, P.C.

; STREET: 1755 S. Jefferson Davis Highway, Suite 400

; CITY: Arlington

; STATE: VA

; COUNTRY: USA

; ZIP: 22202

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: FastSeq Version 1.5

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/612,840A

; FILING DATE: 20-MAR-1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 5-236975

; FILING DATE: 22-SEP-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: No. 5856126man F. Oblon

; REGISTRATION NUMBER: 24,618

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 703-413-3000

; TELEFAX: 703-413-2220

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

```
;
; LENGTH: 690 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Crotalus horridus horridus
; STRAIN:
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 66..512
US-08-612-840A-7

Alignment Scores:
Pred. No.: 4,44e-08 Length: 690
Score: 111.00 Matches: 18
Percent Similarity: 72.41% Conservatives: 3
Best Local Similarity: 62.07% Mismatches: 8
Query Match: 65.29% Indels: 0
DB: 2 Gaps: 0

US-09-938-114-3 (1-29) x US-08-612-840A-7 (1-690)

QY 1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
Db 141 GAATGTCCTCCCGTGGTCTTCTGATGCGGTATTGCTACAGGCCCTTCAAACACAG 200
QY 21 LysThrTrpThrAspAlaGluSerPhe 29
Db 201 ATGACCTGGCCGATGACAGAGGTTTC 227

RESULT 4
US-08-340-428B-1
; Sequence 1, Application US/08340428B
; Patent No. 5648465
; GENERAL INFORMATION:
; APPLICANT: MARGOLIS, Richard U.
; APPLICANT: RAUCH, Uwe
; APPLICANT: MARGOLIS, Renee K.
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES FOR A
; TITLE OF INVENTION: NEUROCAN AS A CHONDROITIN SULFATE PROTEOGLYCAN
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340,428B
; FILING DATE: 14 No. 5648465ember 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/922,911
; FILING DATE: 03 August 1992
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Browdy, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: Margolis=1A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 690 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 66..512
```

```
;
; LENGTH: 5191 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 77..3847
US-08-340-428B-1

Alignment Scores:
Pred. No.: 0.00661 Length: 5191
Score: 84.00 Matches: 11
Percent Similarity: 57.69% Conservatives: 4
Best Local Similarity: 42.31% Mismatches: 11
Query Match: 49.41% Indels: 0
DB: 1 Gaps: 0

US-09-938-114-3 (1-29) x US-08-340-428B-1 (1-5191)

QY 2 CysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSerLys 21
Db 3161 TGGGACCGTGGCTGGCACAATTCAGGCGCCACTGCTACCGCTACTTTGCTCATCGCGG 3220
QY 22 ThrTrpThrAspAlaGlu 27
Db 3221 GCCTGGGAGGACGACGAG 3238

RESULT 5
PCT-US93-07306-1
; Sequence 1, Application PC/TUS9307306
; GENERAL INFORMATION:
; APPLICANT: MARGOLIS, Richard U.
; APPLICANT: RAUCH, Uwe
; APPLICANT: MARGOLIS, Renee K.
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES FOR A
; TITLE OF INVENTION: NEUROCAN AS A CHONDROITIN SULFATE PROTEOGLYCAN
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07306
; FILING DATE: 03-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/922,911
; FILING DATE: 03-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Townsend, Guy K.
; REGISTRATION NUMBER: 34,033
; REFERENCE/DOCKET NUMBER: Margolis=1A PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5191 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 77..3847
```

PCT-US93-07306-1

Alignment Scores:
Pred. No.: 0.00661 Length: 5191
Score: 84.00 Matches: 11
Percent Similarity: 57.69% Conservative: 4
Best Local Similarity: 42.31% Mismatches: 11
Query Match: 49.41% Indels: 0
DB: 5 Gaps: 0

US-09-938-114-3 (1-29) x PCT-US93-07306-1 (1-5191)

QY 2 CysSerSerAspTrpSerSerTyrrGluclyHisCysTyrrLysValPheLysGlnSerLys 21

Db 3161 TGCAGCGTGGCTGGCACAATTCAGGGCCACTGCTACCGCTACTTGTCTCATCGCGGG 3220

QY 22 ThrTrpThrAspAlaGlu 27

Db 3221 GCCTGGGAGGAGCGAG 3238

RESULT 6

PCT-US95-03747-1

Sequence 1, Application PC/TUS9503747

GENERAL INFORMATION:

APPLICANT: LA JOLLA CANCER RESEARCH FOUNDATION

TITLE OF INVENTION: Brevicin, A Glial Cell Proteoglycan

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: Campbell and Flores

STREET: 4370 La Jolla Village Drive, Suite 700

CITY: San Diego

STATE: California

COUNTRY: USA

ZIP: 92122

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/03747

FILING DATE: 27-MAR-1995

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Imbra, Richard J.

REGISTRATION NUMBER: 37,643

REFERENCE/DOCKET NUMBER: PP-LJ 1453

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 535-9001

TELEFAX: (619) 535-8949

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 3259 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

FEATURE:

NAME/KEY: CDS

LOCATION: 112..2848

PCT-US95-03747-1

Alignment Scores:
Pred. No.: 0.0214 Length: 3259
Score: 79.00 Matches: 11
Percent Similarity: 66.67% Conservative: 7
Best Local Similarity: 40.74% Mismatches: 9
Query Match: 46.47% Indels: 0
DB: 5 Gaps: 0

US-09-938-114-3 (1-29) x PCT-US95-03747-1 (1-3259)

QY 2 CysSerSerAspTrpSerSerTyrrGluclyHisCysTyrrLysValPheLysGlnSerLys 21

Db 3161 TGCAGCGTGGCTGGCACAATTCAGGGCCACTGCTACCGCTACTTGTCTCATCGCGGG 3220

Db 2176 TGCAGCCCCGGTTGGGAGCGCTTCCAGGTCCTGTCTACAGCACTTTTTCGCCGAGG 2235

QY 22 ThrTrpThrAspAlaGluSer 28

Db 2236 AGCTGGGAGGAGCGCGAGAAC 2256

RESULT 7

5180808-1

Patent No. 5180808

APPLICANT: RUOSLAHTI, ERKKI I.

TITLE OF INVENTION: VERSICAN CORE PROTEIN, NUCLEIC ACID

SEQUENCES ENCODING THE SAME, NUCLEIC ACID PROBES, ANTI-VERSICAN

ANTIBODIES, AND METHODS OF DETECTING THE SAME

NUMBER OF SEQUENCES: 4

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/441,179

FILING DATE: 27-NOV-1989

SEQ ID NO: 1:

LENGTH: 8224

5180808-1

Alignment Scores:

Pred. No.: 0.27

Score: 75.00 Length: 8224

Percent Similarity: 53.85% Matches: 11

Best Local Similarity: 42.31% Conservative: 3

Query Match: 44.12% Mismatches: 12

DB: 6 Indels: 0

Gaps: 0

US-09-938-114-3 (1-29) x 5180808-1 (1-8224)

QY 2 CysSerSerAspTrpSerSerTyrrGluclyHisCysTyrrLysValPheLysGlnSerLys 21

Db 6810 TGTGACTATGGCTGGCACAATTCAGGGCGAGTGACAAATACTTTGCCATCGAGCG 6869

QY 22 ThrTrpThrAspAlaGlu 27

Db 6870 ACATGGGATGCAGCTGAA 6887

RESULT 8

US-08-543-246B-10

Sequence 10, Application US/08543246B

Patent No. 6262244

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: DNA and amino acid sequence specific for

natural killer cells

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESSEE: Michael W. Glynn

ADDRESSEE: No. 6262244artis Corporation

STREET: 564 Morris Avenue

CITY: Summit,

STATE: NJ

COUNTRY: US

ZIP: 07901-1027

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/543,246B

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/676,663

FILING DATE: 28-MAR-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US92/02469

FILING DATE: 27-MAR-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/122,514

FILING DATE: 24-SEP-1993

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Kassenoff, Melvyn M.
; REGISTRATION NUMBER: 26,389
; REFERENCE/DOCKET NUMBER: 118-7704/PCT/CONT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 908-522-6927
; TELEFAX: 908-522-6955
; INFORMATION FOR SEQ ID NO: 10:
; LENGTH: 402 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-543-246B-10
;
Alignment Scores:
Pred. No.: 0.0132 Length: 402
Score: 73.00 Matches: 9
Percent Similarity: 66.67% Conservative: 9
Best Local Similarity: 33.33% Mismatches: 9
Query Match: 42.94% Indels: 0
DB: 3 Gaps: 0
;
US-09-938-114-3 (1-29) x US-08-543-246B-10 (1-402)
Qy 2 CysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSerLys 21
Db 49 TGTCCTAAAACTGGATATGTTACAAAATAACTGCTACCAATTTTGTGATGAGAGTAAA 108
Qy 22 ThrTrpThrAspAlaGluSer 28
Db 109 AACTGGTATGAGAGCCAGGCT 129
;
RESULT 9
US-08-543-246B-14
; Sequence 14, Application US/08543246B
; Patent No. 6262244
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: DNA and amino acid sequence specific for
; TITLE OF INVENTION: natural killer cells
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Michael W. Glynn
; ADDRESSEE: No. 6262244artis Corporation
; STREET: 564 Morris Avenue
; CITY: Summit,
; STATE: NJ
; COUNTRY: US
; ZIP: 07901-1027
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/543,246B
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/676,663
; FILING DATE: 28-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/02469
; FILING DATE: 27-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/122,514
; FILING DATE: 24-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Kassenoff, Melvyn M.
; REGISTRATION NUMBER: 26,389
```

```
; REFERENCE/DOCKET NUMBER: 118-7704/PCT/CONT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 908-522-6927
; TELEFAX: 908-522-6955
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 648 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-543-246B-14
;
Alignment Scores:
Pred. No.: 0.0238 Length: 648
Score: 73.00 Matches: 9
Percent Similarity: 66.67% Conservative: 9
Best Local Similarity: 33.33% Mismatches: 9
Query Match: 42.94% Indels: 0
DB: 3 Gaps: 0
;
US-09-938-114-3 (1-29) x US-08-543-246B-14 (1-648)
Qy 2 CysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSerLys 21
Db 295 TGTCCTAAAACTGGATATGTTACAAAATAACTGCTACCAATTTTGTGATGAGAGTAAA 354
Qy 22 ThrTrpThrAspAlaGluSer 28
Db 355 AACTGGTATGAGAGCCAGGCT 375
;
RESULT 10
US-09-055-095-2
; Sequence 2, Application US/09055095
; Patent No. 5945308
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Patterson, Chandra
; APPLICANT: Corley, Neil C.
; APPLICANT: Sacher, Susan
; TITLE OF INVENTION: HUMAN OXIDIZED LDL RECEPTOR
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/055,095
; FILING DATE: Filed Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0500 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1740 base pairs
```

```
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: LUNGNOT09
; CLONE: 1355922
US-09-055-095-2

Alignment Scores:
Pred. No.: 0.0803 Length: 1740
Score: 73.00 Matches: 10
Percent Similarity: 60.71% Conservative: 7
Best Local Similarity: 35.71% Mismatches: 11
Query Match: 42.94% Indels: 0
DB: 2 Gaps: 0

US-09-938-114-3 (1-29) x US-09-055-095-2 (1-1740)
Qy 2 CysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSerIys 21
Db 446 TGTACAGACAAATGGAATGCGTAGAGACAAATGCTACCAAGTTCTATAAAGACAGCAAA 505
Qy 22 ThrTrpThrAspAlaGluSerPhe 29
Db 506 AGTTGGGAGGACTGTAATATTTC 529

RESULT 11
US-08-543-246B-8
; Sequence 8, Application US/08543246B
; Patent No. 6262244
; GENERAL INFORMATION:
; TITLE OF INVENTION: DNA and amino acid sequence specific for
; TITLE OF INVENTION: natural killer cells
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Michael W. Glynn
; ADDRESSEE: No. 6262244artis Corporation
; STREET: 564 Morris Avenue
; CITY: Summit,
; STATE: NJ
; COUNTRY: US
; ZIP: 07901-1027
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/543,246B
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/676,663
; FILING DATE: 28-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/02469
; FILING DATE: 27-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/122,514
; FILING DATE: 24-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Kassenoff, Melvyn M.
; REGISTRATION NUMBER: 26,389
; REFERENCE/DOCKET NUMBER: 118-7704/PCT/CONT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 908-522-6927
; TELEFAX: 908-522-6955
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1755 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
```

```
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 339..986
; FEATURE:
; NAME/KEY: mat.peptide
; LOCATION: 339..986
US-08-543-246B-8

Alignment Scores:
Pred. No.: 0.0811 Length: 1755
Score: 73.00 Matches: 9
Percent Similarity: 66.67% Conservative: 9
Best Local Similarity: 33.33% Mismatches: 9
Query Match: 42.94% Indels: 0
DB: 3 Gaps: 0

US-09-938-114-3 (1-29) x US-08-543-246B-8 (1-1755)
Qy 2 CysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSerIys 21
Db 633 TGTCTAAAACTGGATATGTTACAAAAATAACTGCTACCAATTTTGTGAGAGATAAA 692
Qy 22 ThrTrpThrAspAlaGluSer 28
Db 693 AACTGGTATGAGAGCCAGGCT 713

RESULT 12
US-09-996-243-318
; Sequence 318, Application US/09996243
; Patent No. 6478825
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730P1C13
; CURRENT APPLICATION NUMBER: US/09/996,243
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
```

1	PRIOR FILING DATE: 1998-06-17	
2	PRIOR APPLICATION NUMBER: 60/089598	
3	PRIOR FILING DATE: 1998-06-17	
4	PRIOR APPLICATION NUMBER: 60/089599	
5	PRIOR FILING DATE: 1998-06-17	
6	PRIOR APPLICATION NUMBER: 60/089600	
7	PRIOR FILING DATE: 1998-06-17	
8	PRIOR APPLICATION NUMBER: 60/089653	
9	PRIOR FILING DATE: 1998-06-17	
10	PRIOR APPLICATION NUMBER: 60/089801	
11	PRIOR FILING DATE: 1998-06-18	
12	PRIOR APPLICATION NUMBER: 60/089907	
13	PRIOR FILING DATE: 1998-06-18	
14	PRIOR APPLICATION NUMBER: 60/089908	
15	PRIOR FILING DATE: 1998-06-18	
16	PRIOR APPLICATION NUMBER: 60/089947	
17	PRIOR FILING DATE: 1998-06-19	
18	PRIOR APPLICATION NUMBER: 60/089948	
19	PRIOR FILING DATE: 1998-06-19	
20	PRIOR APPLICATION NUMBER: 60/089952	
21	PRIOR FILING DATE: 1998-06-19	
22	PRIOR APPLICATION NUMBER: 60/090246	
23	PRIOR FILING DATE: 1998-06-22	
24	PRIOR APPLICATION NUMBER: 60/090252	
25	PRIOR FILING DATE: 1998-06-22	
26	PRIOR APPLICATION NUMBER: 60/090254	
27	PRIOR FILING DATE: 1998-06-22	
28	PRIOR APPLICATION NUMBER: 60/090349	
29	PRIOR FILING DATE: 1998-06-23	
30	PRIOR APPLICATION NUMBER: 60/090355	
31	PRIOR FILING DATE: 1998-06-23	
32	PRIOR APPLICATION NUMBER: 60/090429	
33	PRIOR FILING DATE: 1998-06-24	
34	PRIOR APPLICATION NUMBER: 60/090431	
35	PRIOR FILING DATE: 1998-06-24	
36	PRIOR APPLICATION NUMBER: 60/090435	
37	PRIOR FILING DATE: 1998-06-24	
38	PRIOR APPLICATION NUMBER: 60/090444	
39	PRIOR FILING DATE: 1998-06-24	
40	PRIOR APPLICATION NUMBER: 60/090445	
41	PRIOR FILING DATE: 1998-06-24	
42	PRIOR APPLICATION NUMBER: 60/090472	
43	PRIOR FILING DATE: 1998-06-24	
44	PRIOR APPLICATION NUMBER: 60/090535	
45	PRIOR FILING DATE: 1998-06-24	
46	PRIOR APPLICATION NUMBER: 60/090557	
47	PRIOR FILING DATE: 1998-06-24	
48	PRIOR APPLICATION NUMBER: 60/090676	
49	PRIOR FILING DATE: 1998-06-25	
50	PRIOR APPLICATION NUMBER: 60/090678	
51	PRIOR FILING DATE: 1998-06-25	
52	PRIOR APPLICATION NUMBER: 60/090690	
53	PRIOR FILING DATE: 1998-06-25	
54	PRIOR APPLICATION NUMBER: 60/090696	
55	PRIOR FILING DATE: 1998-06-25	
56	PRIOR APPLICATION NUMBER: 60/090862	
57	PRIOR FILING DATE: 1998-06-26	
58	PRIOR APPLICATION NUMBER: 60/090863	
59	PRIOR FILING DATE: 1998-06-26	
60	PRIOR APPLICATION NUMBER: 60/091360	
61	PRIOR FILING DATE: 1998-07-01	
62	PRIOR APPLICATION NUMBER: 60/091478	
63	PRIOR FILING DATE: 1998-07-02	
64	PRIOR APPLICATION NUMBER: 60/091544	
65	PRIOR FILING DATE: 1998-07-01	

;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Alignment Scores:
Pred. No.: 0.086 Length: 1841
Score: 73.00 Matches: 10
Percent Similarity: 60.71% Conservatives: 7
Best Local Similarity: 35.71% Mismatches: 11
Query Match: 42.94% Indels: 0
DB: 4 Gaps: 0

US-09-938-114-3 (1-29) x US-09-996-243-318 (1-1841)

Qy 2 CysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSerLys 21
Db 552 TGTACAGCAATGCAATGCGATGGAGACAATTGCTACCAATGCTATATAAGACAGCAAA 611

Qy 22 ThrTrpThrAspAlaGluSerPhe 29

Db 612 AGTGGGAGGACTGTAAATATTC 635

RESULT 13

US-08-840-062-1

; Sequence 1, Application US/08840062

; Patent No. 6117977

; GENERAL INFORMATION:

; APPLICANT: LASKY, LAURENCE A.

; APPLICANT: WU, KAI

; TITLE OF INVENTION: TYPE C LECTINS

; NUMBER OF SEQUENCES: 15

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genentech, Inc.

; STREET: 460 Point San Bruno Blvd

; CITY: South San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94080

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WinPatIn (Genentech)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/840,062

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Dreger, Ginger R.

; REGISTRATION NUMBER: 33,055

; REFERENCE/DOCKET NUMBER: F1019R1

; TELEPHONE: 415/225-3216

; TELEFAX: 415/952-9881

; TELEX: 910/371-7168

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 4588 base pairs

; TYPE: Nucleic Acid

; STRANDEDNESS: Single

; TOPOLOGY: Linear

US-08-840-062-1

Alignment Scores:

Pred. No.: 0.264 Length: 4588
Score: 73.00 Matches: 7
Percent Similarity: 66.67% Conservatives: 11
Best Local Similarity: 25.93% Mismatches: 9
Query Match: 42.94% Indels: 0
DB: 3 Gaps: 0

US-09-938-114-3 (1-29) x US-08-840-062-1 (1-4588)

Qy 1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
Db 1242 GAATGTGACCCAGCTGGCAGCCCTTCACGGGCCACTGTCTACCGCTGCAGCGCGAGAAG 1301

Qy 21 LysThrTrpThrAspAlaGlu 27

Db 1302 CGCAGCTGGCAGGAGTCCAAG 1322

RESULT 14

US-08-840-062-3

; Sequence 3, Application US/08840062

; Patent No. 6117977

; GENERAL INFORMATION:

; APPLICANT: LASKY, LAURENCE A.

; APPLICANT: WU, KAI

; TITLE OF INVENTION: TYPE C LECTINS

; NUMBER OF SEQUENCES: 15

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genentech, Inc.

; STREET: 460 Point San Bruno Blvd

; CITY: South San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94080

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WinPatIn (Genentech)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/840,062

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Dreger, Ginger R.

; REGISTRATION NUMBER: 33,055

; REFERENCE/DOCKET NUMBER: F1019R1

; TELEPHONE: 415/225-3216

; TELEFAX: 415/952-9881

; TELEX: 910/371-7168

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 4771 base pairs

; TYPE: Nucleic Acid

; STRANDEDNESS: Single

; TOPOLOGY: Linear

US-08-840-062-3

Alignment Scores:

Pred. No.: 0.277 Length: 4771
Score: 73.00 Matches: 7
Percent Similarity: 66.67% Conservatives: 11
Best Local Similarity: 25.93% Mismatches: 9
Query Match: 42.94% Indels: 0
DB: 3 Gaps: 0

US-09-938-114-3 (1-29) x US-08-840-062-3 (1-4771)

Qy 1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
Db 1153 GAGTGGAGCGGAGCTGGCAGCCCTTCACGGGCCACTGTCTACCGCTGCAGCGCGAGAAG 1212

Qy 21 LysThrTrpThrAspAlaGlu 27

Db 1213 CGCAGCTGGCAGGAGTCCAAG 1233
RESULT 15
US-08-220-603A-1
; Sequence 1, Application US/08220603A
; Patent No. 5612190
; GENERAL INFORMATION:
; APPLICANT: H. Arita et al.
; TITLE OF INVENTION: A Bovine Group I Phospholipase A2 Receptor
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: c/o FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4 MB format
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS ver. 2.11
; SOFTWARE: Word Perfect (ASCII file)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/220,603A
; FILING DATE: 30-MAR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5-72323
; FILING DATE: 30-March-1993
; APPLICATION NUMBER: JP 5-206686
; FILING DATE: 20-August-1993
; APPLICATION NUMBER: JP 6-40177
; FILING DATE: 10-March-1994
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4978
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE:
; ORGANISM: bovine
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 280 to 4668
; IDENTIFICATION METHOD: by similarity with known sequence or to
; IDENTIFICATION METHOD: an established consensus
US-08-220-603A-1

Alignment Scores:
Pred. No.: 0.414 Length: 4978
Score: 72.00 Matches: 10
Percent Similarity: 64.00% Conservative: 6
Best Local Similarity: 40.00% Mismatches: 9
Query Match: 42.35% Indels: 0
DB: 1 Gaps: 0

US-09-938-114-3 (1-29) x US-08-220-603A-1 (1-4978)

Qy 2 CysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSerLys 21
Db 1411 TGTGAGCTGGCTGGATCCCAACACCGTAATTGCTATTAACCTGCAGAAAGAAAAAAG 1470
Qy 22 ThrTrpThrAspAla 26
Db 1471 ACCTGGATGAGGCT 1485

Search completed: December 8, 2003, 17:27:03
Job time : 25.8686 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 8, 2003, 16:08:50 ; Search time 115.337 Seconds
(without alignments)
835.675 Million cell updates/sec

Title: US-09-938-114-3

Perfect score: 170

Sequence: 1 DCSDDSSVYGHCKYKFKQKWTDAESP 29

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2201672 seqs, 1661799599 residues

Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US0938114/runat_08122003_091002_22963/app_query.fasta_1.725
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blowum62
-TRANS=human40.cdi -LIST=45 -DLOCALIGN=200 -THR SCORE=pct-THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US0938114 @CGN 1 1 504 @runat_08122003_091002_22963
-NCPU=6 -ICPU=3 -NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:**

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:**
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:**
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:**
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:**
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:**
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:**
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:**
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:**
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:**
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:**
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:**
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:**
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:**
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:**
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:**
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:**
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:**
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:**

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description

1	134	78.8	483	10	US-09-929-230-1	Sequence 1, Appli
2	134	78.8	483	13	US-10-226-420-1	Sequence 1, Appli
3	127	74.7	721	10	US-09-929-230-4	Sequence 4, Appli
4	127	74.7	721	13	US-10-226-420-4	Sequence 4, Appli
5	119	70.0	453	10	US-09-929-230-3	Sequence 3, Appli
6	119	70.0	453	13	US-10-226-420-3	Sequence 3, Appli
7	115	67.6	456	10	US-09-929-230-6	Sequence 6, Appli
8	115	67.6	456	13	US-10-226-420-6	Sequence 6, Appli
9	115	67.6	544	11	US-09-938-114-1	Sequence 1, Appli
10	111	65.3	690	10	US-09-969-763-2	Sequence 2, Appli
11	110	64.7	580	10	US-09-929-230-7	Sequence 7, Appli
12	110	64.7	580	13	US-10-226-420-7	Sequence 7, Appli
13	103	60.6	432	10	US-09-929-230-9	Sequence 9, Appli
14	103	60.6	432	13	US-10-226-420-9	Sequence 9, Appli
15	97	57.1	725	10	US-09-929-230-10	Sequence 10, Appl
16	97	57.1	725	13	US-10-226-420-10	Sequence 10, Appl
17	91	53.5	474	10	US-09-929-230-12	Sequence 12, Appl
18	91	53.5	474	13	US-10-226-420-12	Sequence 12, Appl
19	84	49.4	195	13	US-10-029-386-19055	Sequence 19055, A
20	84	49.4	492	9	US-09-801-438-3	Sequence 3, Appli
21	84	49.4	512	13	US-10-029-386-5299	Sequence 5299, Ap
22	84	49.4	6310	13	US-10-241-220-26	Sequence 26, Appl
23	82	48.2	495	10	US-09-893-737-19	Sequence 19, Appl
24	82	48.2	629	9	US-09-801-438-1	Sequence 1, Appli
25	82	48.2	655	10	US-09-965-528-50	Sequence 50, Appl
26	82	48.2	2652	15	US-10-195-970-4	Sequence 4, Appli
27	82	48.2	2652	15	US-10-195-970-5	Sequence 5, Appli
28	82	48.2	3153	11	US-09-759-130B-343	Sequence 343, App
29	82	48.2	3153	13	US-10-188-495-73	Sequence 73, Appl
30	82	48.2	3153	15	US-10-189-123-73	Sequence 73, Appl
31	76	44.7	634	15	US-10-050-704-83	Sequence 83, Appl
32	76	44.7	2694	15	US-10-050-704-40	Sequence 40, Appl
33	76	44.7	3476	13	US-10-237-496-51	Sequence 51, Appl
34	76	44.7	3476	13	US-10-242-074-51	Sequence 51, Appl
35	76	44.7	3476	13	US-10-242-505-51	Sequence 51, Appl
36	76	44.7	3476	13	US-10-242-574-51	Sequence 51, Appl
37	76	44.7	3476	13	US-10-243-261-51	Sequence 51, Appl
38	76	44.7	3476	13	US-10-243-282-51	Sequence 51, Appl
39	76	44.7	3476	13	US-10-243-402-51	Sequence 51, Appl
40	76	44.7	3476	13	US-10-243-431-51	Sequence 51, Appl
41	76	44.7	3476	13	US-10-245-164-51	Sequence 51, Appl
42	76	44.7	3476	13	US-10-244-972-51	Sequence 51, Appl
43	76	44.7	3476	13	US-10-197-942-51	Sequence 51, Appl
44	76	44.7	3476	13	US-10-238-196-51	Sequence 51, Appl
45	76	44.7	3476	13	US-10-245-013-51	Sequence 51, Appl

ALIGNMENTS

RESULT 1
US-09-929-230-1
; Sequence 1, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/09/929,230
; CURRENT FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 483
; TYPE: DNA
; ORGANISM: Sistrurus miliarius
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(455)
; NAME/KEY: misc feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Zsnk2
US-09-929-230-1

```
Alignment Scores:
Pred. No.: 1.24e-12 Length: 483
Score: 134.00 Matches: 22
Percent Similarity: 82.76% Conservative: 2
Best Local Similarity: 75.86% Mismatches: 5
Query Match: 78.82% Indels: 0
DB: 10 Gaps: 0

US-09-938-114-3 (1-29) x US-09-929-230-1 (1-483)
Qy 1 AspCysSerSerAspTrrpSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
Db 60 GATTGTCCTCTGACTGGTCTCTCTATGATCAGCATTGCTACAGGCTTTCAGTGAATC 119

Qy 21 LysThrTrpThrAspAlaGluSerPhe 29
Db 120 AAAACCTGGGATGATGCAGAGAGTTTC 146

RESULT 2
US-10-226-420-1
; Sequence 1, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/10/226,420
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 483
; TYPE: DNA
; ORGANISM: Sistrurus miliarius
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(455)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Zank2
US-10-226-420-1

Alignment Scores:
Pred. No.: 1.24e-12 Length: 483
Score: 134.00 Matches: 22
Percent Similarity: 82.76% Conservative: 2
Best Local Similarity: 75.86% Mismatches: 5
Query Match: 78.82% Indels: 0
DB: 13 Gaps: 0

US-09-938-114-3 (1-29) x US-10-226-420-1 (1-483)
Qy 1 AspCysSerSerAspTrrpSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
Db 60 GATTGTCCTCTGACTGGTCTCTCTATGATCAGCATTGCTACAGGCTTTCAGTGAATC 119

Qy 21 LysThrTrpThrAspAlaGluSerPhe 29
Db 120 AAAACCTGGGATGATGCAGAGAGTTTC 146

RESULT 3
US-09-929-230-4
; Sequence 4, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/09/929,230

Alignment Scores:
Pred. No.: 2.93e-11 Length: 721
Score: 127.00 Matches: 21
Percent Similarity: 79.31% Conservative: 2
Best Local Similarity: 72.41% Mismatches: 6
Query Match: 74.71% Indels: 0
DB: 10 Gaps: 0

US-09-938-114-3 (1-29) x US-09-929-230-4 (1-721)
Qy 1 AspCysSerSerAspTrrpSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
Db 160 GATTGTCCTCTGTTGGTCTCTCTATGATCAGCATTGCTACAGGCTTTCAGTGAATC 219

Qy 21 LysThrTrpThrAspAlaGluSerPhe 29
Db 220 AAGACGTGGGACGATGCAGAGAGTTTC 246

RESULT 4
US-10-226-420-4
; Sequence 4, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/10/226,420
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 721
; TYPE: DNA
; ORGANISM: Sistrurus miliarius
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (91)...(546)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Zank3
US-10-226-420-4

Alignment Scores:
Pred. No.: 2.93e-11 Length: 721
Score: 127.00 Matches: 21
Percent Similarity: 79.31% Conservative: 2
Best Local Similarity: 72.41% Mismatches: 6
Query Match: 74.71% Indels: 0
DB: 13 Gaps: 0

US-09-938-114-3 (1-29) x US-10-226-420-4 (1-721)
Qy 1 AspCysSerSerAspTrrpSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
Db 160 GATTGTCCTCTGTTGGTCTCTCTATGATCAGCATTGCTACAGGCTTTCAGTGAATC 219
```

```

Qy 21 LysThrTrpThrAspAlaGluSerPhe 29
    |||||
Db 220 AAGAGTGGGACGATGCAGAGGTTTC 246

RESULT 5
US-09-929-230-3
; Sequence 3, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/09/929,230
; CURRENT FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 453
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: This degenerate nucleotide sequence encodes the
; NAME/KEY: misc feature
; LOCATION: 9, 12, 18, 21, 24, 27, 30, 36, 39, 42, 45, 48, 51, 54, 57,
; LOCATION: 66, 69, 78, 81, 105, 111, 117, 123, 135, 141, 153, 162,
; LOCATION: 168, 171, 174, 177, 180, 189, 192, 204, 210, 213, 219, 222,
; LOCATION: 225, 231, 234, 243, 246, 258, 261, 270, 297, 306, 312, 324
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: 330, 333, 336, 339, 348, 351, 357, 360, 366, 372, 390, 393,
; LOCATION: 405, 408, 411, 417, 432, 435, 438, 441, 447, 450, 453
; OTHER INFORMATION: n = A,T,C or G
US-10-226-420-3
Alignment Scores:
Pred. No.: 3.33e-10 Length: 453
Score: 119.00 Matches: 18
Percent Similarity: 68.97% Conservative: 2
Best Local Similarity: 62.07% Mismatches: 9
Query Match: 70.00% Indels: 0
DB: 13 Gaps: 0

US-09-938-114-3 (1-29) x US-10-226-420-3 (1-453)
Qy 1 AspCysSerSerAspTrpSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
    |||||
Db 58 GAYTGYCCNWSNGAYTGGWSNWSNTAYGAYCARCAITGYTAYARGTNTTWSNGARYTN 117
    |||||

Qy 21 LysThrTrpThrAspAlaGluSerPhe 29
    |||||
Db 118 AARACNTGGGAYGAYGCGNGARWSNTTY 144
    |||||

RESULT 7
US-09-929-230-6
; Sequence 6, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/09/929,230
; CURRENT FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 456
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: This degenerate nucleotide sequence encodes the
; NAME/KEY: misc feature
; LOCATION: 6, 9, 21, 24, 30, 33, 36, 39, 42, 48, 51, 54, 57, 60, 63,
; LOCATION: 66, 69, 78, 81, 84, 90, 93, 114, 117, 129, 135, 147, 153,
; LOCATION: 162, 171, 177, 180, 186, 189, 192, 201, 204, 213, 216, 222,
; LOCATION: 225, 231, 234, 237, 246, 249, 252, 255, 261, 276, 279
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc feature
; LOCATION: 282, 285, 291, 309, 312, 321, 327, 330, 333, 336, 339, 360,
; LOCATION: 366, 372, 378, 381, 384, 393, 396, 414, 429, 441, 447, 456
; OTHER INFORMATION: n = A,T,C or G
US-09-929-230-6
Alignment Scores:
Pred. No.: 1.53e-09 Length: 456
Score: 115.00 Matches: 18
Percent Similarity: 65.52% Conservative: 1
Best Local Similarity: 62.07% Mismatches: 10
Query Match: 67.65% Indels: 0
DB: 10 Gaps: 0

Qy 1 AspCysSerSerAspTrpSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
    |||||
Db 58 GAYTGYCCNWSNGAYTGGWSNWSNTAYGAYCARCAITGYTAYARGTNTTWSNGARYTN 117
    |||||

Qy 21 LysThrTrpThrAspAlaGluSerPhe 29
    |||||
Db 118 AARACNTGGGAYGAYGCGNGARWSNTTY 144
    |||||

RESULT 6
US-10-226-420-3
; Sequence 3, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/10/226,420
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 453
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: This degenerate nucleotide sequence encodes the

```

```

VENOM OF AGKISTRODON ACUTUS

NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSER: Lyon & Lyon
STREET: 633 West Fifth Street
        Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
        storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSeq for Windows 2.0
CURRENT APPLICATION DATA:
APPLICANT: Lyon & Lyon
APPLICATION NUMBER: US/09/938,114
FILING DATE: 23-Aug-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/058,740
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Chen, Anthony C.
REGISTRATION NUMBER: 38,673
REFERENCE/DOCKET NUMBER: 233/298
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 544 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-938-114-1

Alignment Scores:
Pred. No.: 1,92e-09 Length: 544
Score: 115.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 67.65% Indels: 0
DB: 11 Gaps: 0

US-09-938-114-3 (1-29) x US-09-938-114-1 (1-544)

Qy 10 GluglyHicCysTyrIysValPhelYsGlnSerLystrHrThrAspAlaGluSerPhe 29
Db 1 GRAAGGGCAATGTCTACAAGGCTCTTCAACAACTTAAGACCTGCACAGATGCAGAGAGCTTC 60

RESULT 10
US-09-969-763-2
; Sequence 2, Application US/09969763
; Publication No. US20020198363A1
; GENERAL INFORMATION:
; APPLICANT: FUKUCHI, NAOYUKI
; APPLICANT: KITO, MORIKAZU
; APPLICANT: KAYAHARA, TAKASHI
; APPLICANT: FUTAKI, FUMIE
; APPLICANT: ISHIKAWA, KOHKI
; APPLICANT: SUZUKI, EIICHIRO
; APPLICANT: GONDOH, KEIKO
; APPLICANT: SHIMBA, NOBUHISA
; APPLICANT: YAMADA, NAOYUKI
; TITLE OF INVENTION: PROTEIN HAVING ANTITHROMBOTIC ACTIVITY AND METHOD FOR
; FILE REFERENCE: 2147600S0
; CURRENT APPLICATION NUMBER: US/09/969,763
; CURRENT FILING DATE: 2000-10-25
; PRIOR APPLICATION NUMBER: JP 2000-305279

```

; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2

; LENGTH: 690
; TYPE: DNA

; ORGANISM: Crotalus harridus
; FEATURE:

; NAME/KEY: CDS

; LOCATION: (66)..(512)

; OTHER INFORMATION:

US-09-969-763-2

Alignment Scores:

Pred. No.:	1.18e-08	Length:	690
Score:	111.00	Matches:	18
Percent Similarity:	72.41%	Conservative:	3
Best Local Similarity:	62.07%	Mismatches:	8
Query Match:	65.29%	Indels:	0
DB:	10	Gaps:	0

US-09-938-114-3 (1-29) x US-09-969-763-2 (1-690)

Qy 1 AspCysSerSerAspTyrSerTyrGluclyHisCysTyrLysValPheLysGlnSer 20

Db 141 GAATGTCCTCGGTGGTGGTCTTCCTATGATCGTATGCTACAGCCCTTCAACAAG 200

Qy 21 LysThrTrpThrAspAlaGluSerPhe 29

Db 201 ATGACCTGGCGCGATGCAGAGAGGTTTC 227

RESULT 11

US-09-929-230-7

; Sequence 7, Application US/09929230

; Patent No. US20020161203A1

; GENERAL INFORMATION:

; APPLICANT: Sheppard, Paul O.

; APPLICANT: Bishop, Paul D.

; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS

; FILE REFERENCE: 00-72

; CURRENT APPLICATION NUMBER: US/09/929,230

; CURRENT FILING DATE: 2001-08-13

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 7

; LENGTH: 580

; TYPE: DNA

; ORGANISM: Sistrurus miliarius

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (3)..(434)

; NAME/KEY: misc_feature

; LOCATION: (0)..(0)

; OTHER INFORMATION: Zsnk4

US-09-929-230-7

Alignment Scores:

Pred. No.:	1.38e-08	Length:	580
Score:	110.00	Matches:	17
Percent Similarity:	75.86%	Conservative:	5
Best Local Similarity:	58.62%	Mismatches:	7
Query Match:	64.71%	Indels:	0
DB:	10	Gaps:	0

US-09-938-114-3 (1-29) x US-09-929-230-7 (1-580)

Qy 1 AspCysSerSerAspTyrSerTyrGluclyHisCysTyrLysValPheLysGlnSer 20

Db 36 GATTGTCCTCGATTGGTATGCCATGATGCTATGCTACAGGGTCATCAACAACCTC 95

Qy 21 LysThrTrpThrAspAlaGluSerPhe 29

Db 96 AGGACGTGGCGATGCAGAGAGGTTTC 122

RESULT 12

US-10-226-420-7

; Sequence 7, Application US/10226420

; Publication No. US20030157686A1

; GENERAL INFORMATION:

; APPLICANT: Sheppard, Paul O.

; APPLICANT: Bishop, Paul D.

; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS

; FILE REFERENCE: 00-72

; CURRENT APPLICATION NUMBER: US/10/226,420

; CURRENT FILING DATE: 2002-08-21

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 7

; LENGTH: 580

; TYPE: DNA

; ORGANISM: Sistrurus miliarius

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (3)..(434)

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (0)..(0)

; OTHER INFORMATION: Zsnk4

US-10-226-420-7

Alignment Scores:

Pred. No.:	1.38e-08	Length:	580
Score:	110.00	Matches:	17
Percent Similarity:	75.86%	Conservative:	5
Best Local Similarity:	58.62%	Mismatches:	7
Query Match:	64.71%	Indels:	0
DB:	13	Gaps:	0

US-09-938-114-3 (1-29) x US-10-226-420-7 (1-580)

Qy 1 AspCysSerSerAspTyrSerTyrGluclyHisCysTyrLysValPheLysGlnSer 20

Db 36 GATTGTCCTCGATTGGTATGCCATGATGCTATGCTACAGGGTCATCAACAACCTC 95

Qy 21 LysThrTrpThrAspAlaGluSerPhe 29

Db 96 AGGACGTGGCGATGCAGAGAGGTTTC 122

RESULT 13

US-09-929-230-9

; Sequence 9, Application US/09929230

; Patent No. US20020161203A1

; GENERAL INFORMATION:

; APPLICANT: Sheppard, Paul O.

; APPLICANT: Bishop, Paul D.

; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS

; FILE REFERENCE: 00-72

; CURRENT APPLICATION NUMBER: US/09/929,230

; CURRENT FILING DATE: 2001-08-13

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 9

; LENGTH: 432

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: This degenerate nucleotide sequence encodes the

; OTHER INFORMATION: amino acid sequence of SEQ ID NO:8.

; NAME/KEY: misc_feature

; LOCATION: 6, 15, 18, 21, 24, 27, 42, 45, 57, 78, 81, 93, 96, 99, 111,

; LOCATION: 117, 126, 135, 141, 144, 150, 153, 156, 165, 171, 177, 180,

; LOCATION: 186, 189, 195, 198, 201, 231, 240, 243, 246, 255, 273, 276,

; LOCATION: 285, 291, 294, 297, 300, 315, 327, 330, 345, 348, 360

; OTHER INFORMATION: n = A,T,C or G

; NAME/KEY: misc_feature

; LOCATION: 363, 369, 372, 378, 384, 393, 405, 423, 426

```
; OTHER INFORMATION: n = A,T,C or G
US-09-929-230-9
Alignment Scores:
Pred. No.: 1.34e-07 Length: 432
Score: 103.00 Matches: 16
Percent Similarity: 65.52% Conservative: 3
Best Local Similarity: 55.17% Mismatches: 10
Query Match: 60.59% Indels: 0
DB: 10 Gaps: 0

US-09-938-114-3 (1-29) x US-09-929-230-9 (1-432)
Qy 1 AspCysSerSerAspTrpSerSerTyrGluclyHisCysTyrLysValPheLysGlnSer 20
Db 34 GAYTCYCCNWSNGAYTGTGTAYGAYCARTAYTGTAYMGNGTGNATHAARCARYTN 93
Qy 21 LysThrTrpThrAspAlaGluSerPhe 29
Db 94 MGNACNTGGGAYGAYGCNGARMGNTTY 120

RESULT 14
US-10-226-420-9
; Sequence 9, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/10/226,420
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 432
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: This degenerate nucleotide sequence encodes the
; OTHER INFORMATION: amino acid sequence of SEQ ID NO:8.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 6, 15, 18, 21, 24, 27, 42, 45, 57, 78, 81, 93, 96, 99, 111,
; LOCATION: 117, 126, 135, 141, 144, 150, 153, 156, 165, 171, 177, 180,
; LOCATION: 186, 189, 195, 198, 201, 231, 240, 243, 246, 255, 273, 276,
; LOCATION: 285, 291, 294, 297, 300, 315, 327, 330, 345, 348, 360
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 363, 369, 372, 378, 384, 393, 405, 423, 426
; OTHER INFORMATION: n = A,T,C or G
US-10-226-420-9
Alignment Scores:
Pred. No.: 1.34e-07 Length: 432
Score: 103.00 Matches: 16
Percent Similarity: 65.52% Conservative: 3
Best Local Similarity: 55.17% Mismatches: 10
Query Match: 60.59% Indels: 0
DB: 13 Gaps: 0

US-09-938-114-3 (1-29) x US-10-226-420-9 (1-432)
Qy 1 AspCysSerSerAspTrpSerSerTyrGluclyHisCysTyrLysValPheLysGlnSer 20
Db 34 GAYTCYCCNWSNGAYTGTGTAYGAYCARTAYTGTAYMGNGTGNATHAARCARYTN 93
Qy 21 LysThrTrpThrAspAlaGluSerPhe 29
Db 94 MGNACNTGGGAYGAYGCNGARMGNTTY 120

RESULT 15
```

```
US-09-929-230-10
; Sequence 10, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/09/929,230
; CURRENT FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 725
; TYPE: DNA
; ORGANISM: Sistrurus miliarius
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (88)..(561)
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Zank5
US-09-929-230-10
Alignment Scores:
Pred. No.: 2.52e-06 Length: 725
Score: 97.00 Matches: 15
Percent Similarity: 72.41% Conservative: 6
Best Local Similarity: 51.72% Mismatches: 8
Query Match: 57.06% Indels: 0
DB: 10 Gaps: 0

US-09-938-114-3 (1-29) x US-09-929-230-10 (1-725)
Qy 1 AspCysSerSerAspTrpSerSerTyrGluclyHisCysTyrLysValPheLysGlnSer 20
Db 163 AATTGTCCTCTGTTGGTTGGCTTACGATCAGTATTGCTACAGGGTCATCAACGACTC 222
Qy 21 LysThrTrpThrAspAlaGluSerPhe 29
Db 223 AAGACCTGGGACGATGCAGCGGTTTC 249

Search completed: December 8, 2003, 19:36:05
Job time : 116.337 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:49:54 ; Search time 5.8 Seconds
(without alignments)
480.843 Million cell updates/sec

Title: US-09-938-114-3

Perfect score: 170

Sequence: 1 DCSDDSSSYEGHCYKVFQSKTWDARESF 29.

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 76:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	158	92.9	152	JC7134	agkisacutacin alph
2	133	78.2	146	JC4691	coagulation factor
3	132	77.6	146	JC7135	agkisacutacin beta
4	129	75.9	125	JC5059	bitiscetin beta ch
5	128	75.3	129	JC4329	coagulation factor
6	124	72.9	146	JC7105	aggrstin beta chai
7	123	72.4	152	JC4690	coagulation factor
8	122	71.8	40	S56007	tokaracetin beta c
9	122	71.8	131	JC5058	bitiscetin alpha c
10	119	70.0	125	B47267	botrocetin beta ch
11	117	68.8	133	A47267	botrocetin alpha c
12	115	67.6	40	B56829	alboaggregin-B alp
13	114	67.1	133	JC2415	echicetin beta cha
14	113	66.5	40	S56006	tokaracetin alpha
15	112	65.9	30	A53088	factor IX/factor X
16	112	65.9	123	B42972	coagulation factor
17	104	61.2	40	A56829	alboaggregin-B bet
18	103	60.6	144	PC7027	aggrstin alpha cha
19	100.5	59.1	29	PC4241	multicativase (BC 3
20	97	57.1	30	B53088	factor IX/factor X
21	96	56.5	135	A38609	lectin, galactose-
22	95	55.9	2109	I50421	aggrecan precursor
23	94	55.3	2415	A39086	aggrecan - bovine
24	92	54.1	1340	A39808	proteoglycan core
25	92	54.1	2327	T42630	aggrecan precursor
26	90	52.9	2124	A28452	proteoglycan core
27	90	52.9	2132	A55182	aggrecan precursor
28	89	52.4	71	S55679	flavocetin A - hab
29	84	49.4	1257	S28764	neurocan precursor

30 84 49.4 1268 2 S52781
31 83 48.8 1458 1 A49707
32 82 48.2 883 2 S57653
33 80 47.1 175 2 A37194
34 80 47.1 883 2 S49126
35 79 46.5 912 2 A54423
36 76 44.7 223 2 B46467
37 76 44.7 330 2 T46256
38 76 44.7 3562 2 A47171
39 75 44.1 147 2 A26697
40 75 44.1 862 2 S43922
41 75 44.1 1326 2 B56395
42 75 44.1 1465 2 A56395
43 75 44.1 1643 2 T14274
44 75 44.1 2397 1 A55535
45 75 44.1 2409 1 A60979

ALIGNMENTS

RESULT 1

JC7134

agkisacutacin alpha chain precursor - sharp-nosed viper

N;Alternate names: fibrinogenolytic venom protein

C;Species: Agkistrodon acutus (sharp-nosed viper)

C;Date: 04-Mar-2000 #sequence_revision 04-Mar-2000 #text_change 24-Oct-2000

C;Accession: JC7134; PC7037

R;Cheng, X.; Qian, Y.; Liu, Q.; Li, B.X.Y.; Zhang, M.; Liu, J.

Biochem. Biophys. Res. Commun. 265, 530-535, 1999

A;Title: Purification, characterization, and cDNA cloning of a new fibrinogenolytic venom

A;Reference number: JC7134; MUID:20025379; PMID:10558903

A;Accession: JC7134

A;Molecule type: mRNA

A;Residues: 1-152 <CHE>

A;Cross-references: GB:AF176420

A;Experimental source: venom gland

A;Accession: PC7037

A;Molecule type: protein

A;Residues: 24-53;84-86;87-94;125-136;137-152 <CH2>

C;Superfamily: tetranectin; C-type lectin homology

C;Keywords: disulfide bond; heterodimer; venom

F;1-23/Domain: signal sequence #status predicted <SIG>

F;24-152/Product: agkisacutacin alpha chain #status experimental <MAT>

Query Match 92.9%; Score 158; DB 2; Length 152;
Best Local Similarity 93.1%; Pred. No. 7.9e-13;
Matches 27; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 DCSDDSSSYEGHCYKVFQSKTWDARESF 29

Db 24 DCSGSSSYEGHCYKVFQSKTWADAEF 52

RESULT 2

JC4691

coagulation factor IX/factor X-binding protein chain A precursor - habu

C;Species: Trimeresurus flavoviridis (habu)

C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 20-Jun-2000

C;Accession: JC4691; B39332; JC4330

R;Matsuzaki, R.; Yoshiara, E.; Yamada, M.; Shima, K.; Atoda, H.; Morita, T.

Biochem. Biophys. Res. Commun. 220, 382-387, 1996

A;Title: cDNA cloning of IX/X-BP, a heterogeneous two-chain anticoagulant protein from si

A;Reference number: JC4691; MUID:96184662; PMID:8645314

A;Accession: JC4691

A;Molecule type: mRNA

A;Residues: 1-146 <MAT1>

A;Cross-references: DDBJ:D83332; NID:g1402641; PIDN:BA11888.1; PID:g1402642

A;Experimental source: venom

R;Atoda, H.; Hyuga, M.; Morita, T.

J. Biol. Chem. 266, 14903-14911, 1991

A;Title: The primary structure of coagulation factor IX/factor X-binding protein isolate

tein, tetranectin, and lymphocyte Fc epsilon receptor for immunoglobulin E.

A; Experimental source: snake venom
R; Matsui, T.; Hamako, J.; Suzuki, M.; Hayashi, N.; Ito, M.; Makita, K.; Fujimura, Y.; Oze-
res Commun. Biochem. Cell Mol. Biol. 1, 271-284, 1997

A; Title: Complete amino acid sequence of bitiscetin, a novel von Willebrand factor module
A; Reference number: JC5916
A; Accession: JC5917
A; Molecule type: protein
A; Residues: 1-125 <MA2>

A; Experimental source: venom
C; Comment: This protein is a modulator of a von Willebrand factor modulator.
C; Superfamily: tetranectin; C-type lectin homology
C; Keywords: venom
F; 4-121/Domain: C-type lectin homology <LCH>

Query Match 75.9%; Score 129; DB 2; Length 125;
Best Local Similarity 75.0%; Pred. No. 3e-09;
Matches 21; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 2 CSSDWSSYEGHCYKVFVKOSKTWTDAESF 29
| | | | | | | | | | | | | | | | | | | | | |
Db 4 CLPDWSSYGHCYKVFVKETWADEKF 31

RESULT 5

JC4329
coagulation factor IX-binding protein A chain - habu
C; Species: Trimeresurus flavoviridis (habu)
C; Date: 06-Dec-1995 #sequence_revision 08-Feb-1996 #text_change 13-Mar-1998
C; Accession: JC4329
E; Atoda, H.; Ishikawa, M.; Yoshihara, E.; Sekiya, F.; Morita, T.
J. Biochem. 119, 965-973, 1995
A; Title: Blood coagulation factor IX-binding protein from the venom of Trimeresurus flavo-
A; Reference number: JC4329; PMID: 96318509; PMID: 8749314
A; Accession: JC4329
A; Molecule type: protein
A; Residues: 1-129 <ARO>
C; Comment: This protein binds calcium.
C; Superfamily: tetranectin; C-type lectin homology
C; Keywords: anticoagulant; blood coagulation; calcium binding; venom
F; 2-127/Domain: C-type lectin homology <LCH>
F; 2-13, 30-127, 102-119/Disulfide bonds: #status predicted

Query Match 75.3%; Score 128; DB 2; Length 129;
Best Local Similarity 75.9%; Pred. No. 4.1e-09;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 DCSDWSSEGHGHCYKVFVKOSKTWDASF 29
| | | | | | | | | | | | | | | | | | | | | |
Db 1 DCPGSWSSYEGHCYKPFKYLTWDDAEFR 29

RESULT 6

JC7105
aggretin beta chain - Malayan pit viper
C; Species: Calloselasma rhodostoma (Malayan pit viper)
C; Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 11-May-2000
C; Accession: JC7105
R; Chung, C.H.; Au, L.C.; Huang, T.F.
Biochem. Biophys. Res. Commun. 263, 723-727, 1999
A; Title: Molecular cloning and sequence analysis of aggretin, a collagen-like platelet aggre-
A; Reference number: PC7027; PMID: 99443731; PMID: 10512747
A; Accession: JC7105
A; Molecule type: mRNA
A; Residues: 1-146 <CHU>
A; Experimental source: venom gland
C; Superfamily: tetranectin; C-type lectin homology
C; Keywords: disulfide bond; platelet aggregation; venom

Query Match 72.9%; Score 124; DB 2; Length 146;
Best Local Similarity 69.0%; Pred. No. 1.4e-08;
Matches 20; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 1 DCSDWSSEGHGHCYKVFVKOSKTWTDASF 29

A47267

botrocetin alpha chain - jararaca
N;Alternate names: two chain botrocetin alpha chain
C;Species: Bothrops jararaca (jararaca)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 15-May-1998
C;Accession: A47267; B37958
R;Usami, Y.; Fujimura, Y.; Suzuki, M.; Ozeki, Y.; Nishio, K.; Fukui, H.; Titani, K.
Proc. Natl. Acad. Sci. U.S.A. 90, 928-932, 1993
A;Title: Primary structure of two-chain botrocetin, a von Willebrand factor modulator protein
A;Reference number: A47267; MUID:93157385; PMID:8430107
A;Accession: A47267
A;Molecule type: protein
A;Residues: 1-133 <UGA>
A;Experimental source: venom
A;Note: sequence extracted from NCBI backbone (NCBIP:124085)
R;Fujimura, Y.; Titani, K.; Usami, Y.; Suzuki, M.; Oyama, R.; Matsui, T.; Fukui, H.; Sugimoto, S.
Biochemistry 30, 1957-1964, 1991
A;Title: Isolation and chemical characterization of two structurally and functionally different botrocetin alpha chains
A;Reference number: A37958; MUID:91129280; PMID:1993206
A;Accession: B37958
A;Molecule type: protein
A;Residues: 1-40 <FUU>
C;Complex: heterodimer of alpha and beta (see PIR:B47267) chains
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: hemagglutinin; heterodimer; venom
F;2-128/Domain: C-type lectin homology <LCH>
F;2-13,30-128,103-120/Disulfide bonds: #status experimental
F;80/Disulfide bonds: interchain (co beta-75) #status experimental

Query Match 68.8%; Score 117; DB 2; Length 133;
Best Local Similarity 65.5%; Pred. No. 1e-07;
Matches 19; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 1 DCSSDWSYEGHCYKVKFQSKTWTDAESF 29
||| ||||| : ||| : ||| :
Db 1 DCPGWSYEGHCYKVKFQSKTWTDAERF 29

RESULT 12

B56829
alboaggregin-B alpha chain - green pit viper (fragment)
C;Species: Trimeresurus albolabris (green pit viper)
C;Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 03-May-1996
C;Accession: B56829
R;Yoshida, E.; Fujimura, Y.; Miura, S.; Sugimoto, M.; Fukui, H.; Narita, N.; Usami, Y.;
Biochem. Biophys. Res. Commun. 191, 1386-1392, 1993
A;Title: Alboaggregin-B and botrocetin, two snake venom proteins with highly homologous
A;Reference number: B56829; MUID:93221514; PMID:8466514
A;Accession: B56829
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-40 <YOS>
A;Experimental source: venom
A;Note: sequence extracted from NCBI backbone (NCBIP:128751)
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: disulfide bond; heterodimer

Query Match 67.6%; Score 115; DB 2; Length 40;
Best Local Similarity 62.1%; Pred. No. 6.4e-08;
Matches 18; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 1 DCSSDWSYEGHCYKVKFQSKTWTDAESF 29
||| ||||| : ||| : ||| :
Db 1 DCPGWSYEGHCYKVKFQSKTWTDAERF 29

RESULT 13

JC2415
echicetin beta chain - saw-scaled viper
C;Species: Echis carinatus (saw-scaled viper)
C;Date: 21-Mar-1995 #sequence_revision 26-May-1995 #text_change 15-Oct-1996
C;Accession: JC2415
R;Peng, M.; Holt, J.C.; Niewiarowski, S.

Biochem. Biophys. Res. Commun. 205, 68-72, 1994
A;Title: Isolation, characterization and amino acid sequence of echicetin beta subunit,
A;Reference number: JC2415; MUID:95091801; PMID:7999097
A;Accession: JC2415
A;Molecule type: protein
A;Residues: 1-123 <PEN>
A;Experimental source: venom
C;Comment: This protein inhibits agglutination of fixed platelets induced by several platelet
aggregating factors and also agglutinates.
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: anticoagulant; dimer
F;2-119/Domain: C-type lectin homology <LCH>
F;2-13,30-119,96-111/Disulfide bonds: #status predicted

Query Match 67.1%; Score 114; DB 2; Length 123;
Best Local Similarity 62.1%; Pred. No. 2.3e-07;
Matches 18; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 1 DCSSDWSYEGHCYKVKFQSKTWTDAESF 29
||| ||||| : ||| : ||| :
Db 1 NCLPDWSYEGHCYKVKFQSKTWTDAERF 29

RESULT 14

S56006
tokaracetin alpha chain - Trimeresurus tokarensis (fragment)
N;Alternate names: platelet aggregation inhibitor; platelet antagonist
C;Species: Trimeresurus tokarensis
C;Date: 10-Oct-1995 #sequence_revision 08-Nov-1996 #text_change 07-May-1999
C;Accession: S56006
R;Kawasaki, T.; Taniuchi, Y.; Hisamichi, N.; Fujimura, Y.; Suzuki, M.; Titani, K.; Sakai,
Biochem. J. 308, 947-953, 1995
A;Title: Tokaracetin, a new platelet antagonist that binds to platelet glycoprotein Ib
A;Reference number: S56006; MUID:97104297; PMID:8948455
A;Accession: S56006
A;Molecule type: protein
A;Residues: 1-40 <KAW>
C;Superfamily: tetranectin; C-type lectin homology

Query Match 66.5%; Score 113; DB 2; Length 40;
Best Local Similarity 65.5%; Pred. No. 1.1e-07;
Matches 19; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 1 DCSSDWSYEGHCYKVKFQSKTWTDAESF 29
||| ||||| : ||| : ||| :
Db 1 DCPGWSYEGHCYKVKFQSKTWTDAERF 29

RESULT 15

A53088
factor IX/factor X-binding anticoagulant protein B chain - jararaca (fragment)
C;Species: Bothrops jararaca (jararaca)
C;Date: 13-Sep-1994 #sequence_revision 18-Nov-1994 #text_change 15-Jun-1996
C;Accession: A53088
R;Sekiya, F.; Atoda, H.; Morita, T.
Biochemistry 32, 6892-6897, 1993
A;Title: Isolation and characterization of an anticoagulant protein homologous to botrocetin
A;Reference number: A53088; MUID:9326575; PMID:8334120
A;Accession: A53088
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-30 <SEK>
A;Experimental source: venom
A;Note: sequence extracted from NCBI backbone (NCBIP:135337)

Query Match 65.9%; Score 112; DB 2; Length 30;
Best Local Similarity 62.1%; Pred. No. 1.2e-07;
Matches 18; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 1 DCSSDWSYEGHCYKVKFQSKTWTDAESF 29
||| ||||| : ||| : ||| :
Db 1 DCPGWSYEGHCYKVKFQSKTWTDAERF 29